The Astrophysical Journal, 721:1142–1157, 2010 October 1 doi:10.1088/0004-637X/721/2/1142

CORRECTING VELOCITY DISPERSIONS OF DWARF SPHEROIDAL GALAXIES FOR BINARY ORBITAL MOTION

Quinn E. Minor[1], Greg Martinez[1], James Bullock[1], Manoj Kaplinghat[1], and Ryan Trainor[2]
[1] Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
[2] Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
Received 2010 March 8; accepted 2010 June 4; published 2010 September 8

ABSTRACT

We show that the measured velocity dispersions of dwarf spheroidal galaxies from about 4 to 10 km s^{-1} are unlikely to be inflated by more than 30% due to the orbital motion of binary stars and demonstrate that the intrinsic velocity dispersions can be determined to within a few percent accuracy using two-epoch observations with 1–2 yr as the optimal time interval. The crucial observable is the threshold fraction—the fraction of stars that show velocity changes larger than a given threshold between measurements. The threshold fraction is tightly correlated with the dispersion introduced by binaries, independent of the underlying binary fraction and distribution of orbital parameters. We outline a simple procedure to correct the velocity dispersion to within a few percent accuracy by using the threshold fraction and provide fitting functions for this method. We also develop a methodology for constraining properties of binary populations from both single- and two-epoch velocity measurements by including the binary velocity distribution in a Bayesian analysis.

Key words: binaries: spectroscopic – galaxies: kinematics and dynamics

1. INTRODUCTION

In recent years, a large number of dwarf spheroidal satellite galaxies of the Milky Way have been discovered using the Sloan Digital Sky Survey (Willman et al. 2005; Zucker et al. 2006). These galaxies are much fainter than previously known Milky Way satellites, having larger mass-to-light ratios and velocity dispersions that range from 7.6 km s^{-1} down to 3.3 km s^{-1} (Simon & Geha 2007). Estimating the amount of dark matter contained in these galaxies becomes more susceptible to error than in the larger dwarf spheroidals. This is not only because of the statistical error associated with small stellar samples, but also because sources of velocity contamination constitute a larger fractional error due to the galaxy's small intrinsic velocity dispersion. Potential sources of contamination come from foreground Milky Way stars, atmospheric jitter in red giant stars (Pryor et al. 1988), and an inflated velocity dispersion due to the orbital motion of binary stars (Olszewski et al. 1996). Among these, binaries have been the most difficult to correct for because the binary distribution of velocities in environments beyond the solar neighborhood is not well known and difficult to constrain without a large number of high-precision radial velocity measurements.

The most prominent signature of binary stars in these galaxies is a high-velocity "tail" in the velocity distribution due to short-period binaries, which gives the distribution a higher kurtosis than that expected from a Gaussian. Because of this departure from Gaussianity, the intrinsic velocity dispersion is usually estimated by using a robust estimator such as the biweight (Beers et al. 1990), by discarding velocity outlier stars from the data sample, or by a combination of both approaches (cf. Mateo et al. 1991). While these techniques eliminate the largest component of the binary dispersion, binaries inflate the first-order (Gaussian) component of the dispersion as well. In previous studies, Olszewski et al. (1996) and Hargreaves et al. (1996) used Monte Carlo simulations to show that the dispersion estimated by these techniques is inflated due to binaries by an amount which is small compared to the statistical error in their data sets. Since that time, the number of stars in then-known dwarf spheroidals with spectroscopic data has increased considerably, from less than 100 at the time of Olszewski et al. (1996) to more than 1000 for Draco, Fornax, and Carina (Walker et al. 2009a). For such large samples, we will show that the bias in the first-order dispersion due to binaries is larger than the statistical error. More importantly, for galaxies whose intrinsic dispersion is small, the first-order dispersion may be inflated by somewhat more than the statistical error even in data sets as small as 100 stars.

To improve previous estimates of the intrinsic dispersion for these cases, it is necessary to model the velocity distribution of the binary population and investigate the behavior of the binary dispersion as model parameters such as binary fraction are varied. With this approach it is desirable to find the best possible constraints on the binary velocity distribution and observations taken at multiple epochs are very useful for this purpose. In addition, estimating the binary fraction and other binary properties in galaxies and clusters beyond the solar neighborhood is useful in its own right, since this has been difficult to predict accurately in simulations (Tohline 2002; Goodwin et al. 2007). Binaries also affect higher-order moments of the velocity distribution, e.g., kurtosis, which has been shown in principle to provide useful information about the mass profile of galaxies (Łokas et al. 2005).

In Section 2, we derive an analytic formula for the center-of-mass velocity distribution of binary stars. In Section 3.1, we use this distribution to derive likelihoods for both single- and two-epoch velocity data. In Section 3.2, we define an observable quantity, the threshold fraction, which will prove useful in later sections for characterizing a binary population and correcting the velocity dispersion. We demonstrate in Section 4 how properties of a binary population can be constrained from multi-epoch data. Figure 4 shows that repeat measurements with a baseline of 1–2 yr are sufficient to constrain the binary population. In Section 5, we prove that over a range of velocities of a few km s^{-1}, the binary fraction is nearly degenerate with the parameters characterizing the distribution of periods; we

then use this fact to develop a model-independent method for correcting the dispersion due to binaries by using multi-epoch data. Figure 11 encapsulates our main result, from which we conclude that the velocity dispersions of dwarf spheroidals are unlikely to be inflated by more than 30% by binaries (Section 5.6). The general procedure to correct the velocity dispersion for binaries is summarized in Section 5.5. In Section 6, we outline a method to combine single- and multi-epoch data in a Bayesian analysis and discuss the issue of foreground contamination and how this affects the apparent binary population. Finally, in Section 7, we further explore the degeneracy of model parameters and discuss prospects for constraining the period distribution of binary systems from multi-epoch velocity data.

2. DISTRIBUTION OF VELOCITIES IN THE CENTER-OF-MASS FRAME OF BINARY SYSTEMS

Spectroscopic velocity measurements of binary stars are typically dominated by the primary star due to the difference in luminosities and spectra between the two stars. This is especially true for stars that lie on the red giant branch, for which the luminosity–mass relation steepens drastically. By way of comparison, Duquennoy & Mayor (1991) found the ratio of secondary-to-primary masses of binary stars within the solar neighborhood to peak at $m/M \approx 0.23$ with only ≈20% of the stars having mass ratios larger than 0.6. For systems with a mass ratio of 0.6 and the primary lying near the base of the red giant branch, the luminosity ratio would be at most $l \approx (m/M)^{2.3} \approx 0.3$; this ratio becomes much smaller if the primary star is further along on the red giant branch. Thus, although the cross-correlation function of the stellar spectra may be double-peaked (Tonry & Davis 1979), the spectrum of the primary star will likely dominate the signal unless the mass ratio is close to 1, in which case the velocities of the two stars will be nearly equal. In view of the preceding arguments, for the remainder of this paper we will use the velocity of the primary star in modeling the spectroscopic velocities of binary systems.

To model the velocity distribution of the primary star in a population of binary systems, first we must find its velocity distribution in the center-of-mass frame of the binary. The motion of two stars orbiting each other can be simply expressed in terms of four parameters: the semimajor axis a, eccentricity e, period P, and orbital angle ϕ with respect to the center of mass. If we also specify the orientation of the orbital plane in terms of Euler angles, with the z-axis pointing along the line of sight, the line-of-sight velocity of the primary star in the center-of-mass frame is given by

$$v_z = \frac{2\pi a}{P}(1-e^2)^{-\frac{1}{2}} g_e(\theta, \psi, \phi), \quad (1)$$

$$g_e(\theta, \psi, \phi) \equiv \sin\theta[\cos(\psi - \phi) + e\cos(\phi)], \quad (2)$$

where a is the semimajor axis of the primary star's orbit, and θ and ψ are the second and third Euler angles, respectively. The azimuthal Euler angle does not appear explicitly because v_z is invariant under rotations about the z-axis.

Now taking the log of Equation (1) and substituting Kepler's third law, we find

$$\log P = k - 3\log|v_z|, \quad (3)$$

$$k \equiv \log\left\{\left(\frac{|g_e(\theta, \psi, \phi)|}{\sqrt{1-e^2}}\right)^3 \frac{(2\pi q)^3 m}{(1+q)^2}\right\}, \quad (4)$$

where P is in years, m is the mass of the primary star in solar masses, q is the ratio of secondary-to-primary mass, and v_z is in units of AU yr^{-1}.

As an aside, we note that Equations (3) and (4) can be used to make a back-of-the-envelope estimate of the velocity scale associated with a given orbital period. By averaging over orientations, mass ratio and eccentricity (whose PDF's are given below), we find

$$|v_z| \approx (5.7 \text{ km s}^{-1})\left(\frac{M/M_\odot}{P/\text{yr}}\right)^{\frac{1}{3}}, \quad (5)$$

where M is the mass of the primary star. In dwarf spheroidal galaxies and globular clusters where red giants have masses $M \approx 0.8\,M_\odot$, the above estimate shows that periods longer than a few decades will yield velocities less than 2 km s^{-1}. Meanwhile, velocities larger than 10 km s^{-1} will be dominated by binaries with periods shorter than 1 month. The above estimate is somewhat sensitive to the mass ratio; whereas we used the approximate mean value $q \approx 0.4$ in deriving Equation (5), if a system has a mass ratio $q \approx 0.8$, the coefficient in front becomes ≈10 km s^{-1}.

To find a distribution in the center-of-mass velocity v_z, the distribution of orbital periods must be averaged over all the parameters in Equation (4). We must also average over the time taken to traverse one orbital cycle, with all times being weighted equally. We express the orbital angle ϕ in terms of the eccentric anomaly parameter η, so that a uniform distribution in time corresponds to a distribution $f(\eta) = (1 - e\cos\eta)/2\pi$.

The orbital periods of G-dwarf stars in the solar neighborhood were found by Duquennoy & Mayor (1991) to follow a log-normal distribution with a mean period of 180 yr. Fischer & Marcy (1992) also found a log-normal period distribution for M-dwarfs in the solar neighborhood with a mean period similar to that of the G-dwarfs. In terms of logarithm of the period P, the distribution found by Duquennoy & Mayor (1991) has mean $\mu_{\log P} = 2.24$ and dispersion $\sigma_{\log P} = 2.3$ (where P is in years, and the logarithm is base 10). We shall use this as the fiducial binary model in this paper, but will also allow $\mu_{\log P}$ and $\sigma_{\log P}$ to take on other values.

As with the period distribution, we use distributions of the mass ratio q and eccentricity e observed in G-dwarf stars in the solar neighborhood (Duquennoy & Mayor 1991). When comparing different empirically derived mass functions, Duquennoy & Mayor (1991) showed that the distribution of mass ratios q was best fit by the Gaussian mass function considered in Kroupa et al. (1990) with mean $\bar{q} = 0.23$ and dispersion $\sigma_q = 0.32$. This is somewhat misleading, since for $q > 0.5$ the distribution is in fact consistent with the power-law initial mass functions of Salpeter and Kroupa (Salpeter 1955; Kroupa 2001). At smaller mass ratios the distribution decreases sharply compared to the well-established Kroupa initial mass function of single stars, which implies that small mass ratios are strongly affected by interaction between primary and secondary during the formation of the binary system.

To further complicate matters, Mazeh et al. (1992) found that short-period binaries tend to have higher mass ratios than those found by Duquennoy & Mayor (1991). By analyzing spectroscopic binaries, they found that binaries with periods shorter than 3000 days have mass ratios consistent with a uniform distribution, although the Poisson errors in their analysis were quite large. Subsequent studies (Goldberg et al. 2003; Halbwachs et al. 2004) have shown that the mass ratio distribution in short-period binaries is bimodal (also seen in previous

samples; cf. Trimble 1990), with a peak at low mass ratios similar to that of the long-period binaries but with another peak near $q \approx 1$. In the sample analyzed by Goldberg et al. (2003), the peak at high mass ratios is smaller for primaries with masses larger than 0.6 $M_\odot$, but larger for halo stars. In view of lingering uncertainties in the nature of the mass ratio distribution, for simplicity we will adopt the Gaussian distribution from Duquennoy & Mayor (1991) for long-period binaries ($P >$ 3000 days) and a uniform distribution for short-period binaries ($P <$ 3000 days). For the distribution of the primary mass $f(m)$ we will use the Kroupa initial mass function corrected for binaries (Kroupa 2002).

The distribution of eccentricities $f(e|\log P)$ was found by Duquennoy & Mayor (1991) to have three different regimes depending on the period. For periods of 11 days or shorter, the orbits are circularized due to tidal forces and are therefore approximated to have $e = 0$. For periods between 11 days and 1000 days, $f(e)$ can be approximated by a Gaussian with mean $\bar{e} = 0.25$ and dispersion $\sigma_e = 0.12$. For periods longer than 1000 days, higher eccentricities are more prevalent and the distribution approximately follows $f(e) = \frac{3}{2}e^{1/2}$. Among these, the Gaussian regime (11 days $< P <$ 1000 days) has the greatest impact on the velocity distribution at velocities of order km s^{-1}. While the adopted distributions of mass ratios and eccentricities are undoubtedly only a rough approximation to the true distributions, our central results (presented in Section 5) will prove to be quite insensitive to the nature of the adopted distributions.

An important effect that must be taken into account is the effect of mass transfer between the stars if the primary star is a red giant whose size is larger than the radius of its Roche lobe (Paczyński 1971). In such a case, matter from the surface of the giant will accrete onto the smaller star and the separation between the stars will decrease. The end result is that either the other star will explode in a supernova Ia, or the stars will eventually merge. This effect is not included in the distributions of Duquennoy & Mayor (1991) because their sample consisted entirely of dwarf stars. Therefore, we make the approximation of excluding systems whose primary star is larger than its Roche lobe, assuming the binary to be destroyed over a timescale much less than 1 Gyr. We use an approximation to the radius $r_L(a, q)$ of the Roche lobe given by Eggleton (1983). While this radius is not exactly correct for eccentric orbits, a recent smoothed-particle hydrodynamics simulation of an eccentric binary of mass ratio $q = 0.6$ by Church et al. (2009) found the Roche lobe radius to decrease only slightly with eccentricity. They also derive a fitting function for the Roche lobe radius with an eccentricity e, given by $r_L(e) = r_L(e = 0)(1 - 0.16e)$. We find using this formula that the velocity distribution changes only by a small amount (less than 2% at 5 km s^{-1}) compared to when using the Roche lobe radius evaluated at pericenter, given by $r_L(e = 0)$ above. For the following calculations, we therefore adopt the Roche lobe evaluated at pericenter for an eccentric orbit.

The radius of each star is found by estimating its effective temperature from an isochrone of given age t_g in the stellar population synthesis model of Girardi et al. (2004). This, together with its magnitude, provides an estimate of the stellar radius. We denote $M_V(m; t_g)$ and $R(m; t_g)$ as the absolute V-band magnitude and radius (respectively) of a star of mass m assigned by an isochrone of age t_g. If the star lies on the horizontal branch or asymptotic giant branch, instead of using its present radius (which may be small) we compare its Roche lobe to the largest radius previously attained by the star at the end of its red giant phase.

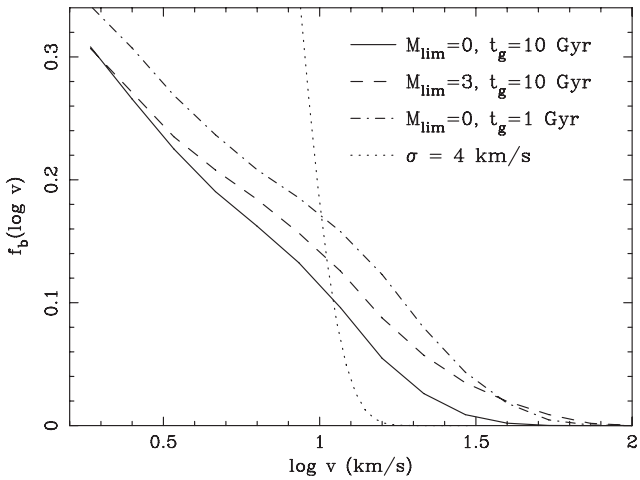


Figure 1. Distribution of velocities in the center-of-mass frame of binary systems, plotted for different absolute magnitude limits $M_{\rm lim} = 0$, 3 and stellar ages $t_g = 1$ Gyr, 10 Gyr. For comparison we also plot a Gaussian velocity distribution (dotted line) in log space with dispersion of 4 km s^{-1} and no binaries. The suppression of binaries due to Roche-lobe overflow becomes important for velocities $\gtrsim$10 km s^{-1}. Except at the turnover point $v \approx 10$ km s^{-1}, the distribution behaves locally as a log-normal to good approximation.

We shall express our formula in terms of an absolute V-band upper magnitude limit $M_{\rm lim}$ and age t_g. Assuming the lower magnitude limit to be near the tip of the red giant branch, we find that the velocity distribution is quite insensitive to the exact value of the lower magnitude limit because the suppression of binaries due to Roche lobe overflow dominates the high-luminosity end of the red giant branch.

Averaging over distributions for all the model parameters and dropping the z subscript for readability, we obtain

$$\begin{aligned} f_b(\log|v|; M_{\rm lim}, t_g) = \frac{3}{8\pi^2}\int_{-1}^{1} d(\cos\theta)\int_0^{2\pi} d\psi \\ \times \int_0^1 f(e|\log P)de \int_0^{2\pi} f(\eta)d\eta \\ \times \int_0^1 f(q|\log P)dq \int_0^{\infty} f(m)dm \cdot \Theta \\ \times [M_{\rm lim} - M_V(m; t_g)] \\ \times \Theta[r_L(a, q) - R(m; t_g)] \\ \times \frac{\exp\left\{\frac{-[3\log|v| - k + \mu_{\log P}]^2}{2\sigma^2_{\log P}}\right\}}{\sqrt{2\pi\sigma^2_{\log P}}}, \end{aligned} \quad (6)$$

where $\Theta[x]$ is the Heaviside step function. The variable k is a function of all the other parameters according to Equation (4), with $\phi \equiv \phi(\eta)$. We use a Monte Carlo simulation to perform the integration over a grid of $\log|v|$ values and interpolate to find $f_b(\log|v|)$. Figure 1 plots the distribution $f_b(\log|v|)$ for different ages and absolute magnitude limits. For comparison, we also plot a Gaussian velocity distribution (dotted line) with a dispersion of 4 km s^{-1} and no binaries. This figure shows that for velocities $\gtrsim$10 km s^{-1}, suppression of binaries due to Roche-lobe overflow becomes evident as a turnover in the velocity distribution.

Asymptotically for large velocities, the velocity distribution in Equation (6) behaves as a log-normal with dispersion similar

to that of the period distribution, $\sigma_{\log P}$. This can be seen as follows: first, at large velocities, the log-normal in the integrand is far from its maximum and therefore varies slowly in the model parameters (Equation (4)) compared to their respective probability distributions, provided k is not large and negative. Therefore as a first approximation, we can apply the method of steepest descents and find the resulting distribution to be log-normal with a mean given by $\log|\bar{v}| = \frac{1}{3}(\bar{k} - \mu_{\log P})$. However, the approximation of slowly varying log-normal is not strictly true, since k becomes large and negative if the mass ratio q and direction function $g_e(\theta, \psi, \phi)$ are close to zero. This means that the mean $\bar{k}$ is in fact a function of v. We find, however, that locally $\bar{k}$ is linear in $\log|v|$ to good approximation, with the result that the velocity distribution still behaves locally as a log-normal but with a somewhat different dispersion from $\sigma_{\log P}$. The local log-normal behavior of the binary velocity distribution has important implications for determining the velocity dispersion due to binaries, which we discuss further in Section 5.

3. VELOCITY DISTRIBUTION OF A POPULATION OF BINARY STARS

3.1. Single- and Two-epoch Likelihood Functions

Suppose that among a population of stars, a fraction B of them are in binary systems. Further suppose that the velocity distribution for stars not in binary systems is Gaussian with dispersion σ_0 and systemic velocity $\bar{v}$. The velocity likelihood function will have the following form:

$$L(v|B, \sigma_0, \bar{v}) = (1 - B)\frac{e^{-(v-\bar{v})^2/2\sigma_0^2}}{\sqrt{2\pi\sigma_0^2}} + BL_b(v|\sigma_0, \bar{v}), \quad (7)$$

where $L_b(v|\sigma_0, \bar{v})$ is the likelihood for binary stars. To derive the binary likelihood, we note that the component of the velocity *not* due to the binary orbit is $v_{nb} = v - v'$, where v' is the line-of-sight component of the velocity in the center-of-mass frame of the binary system. To find the binary likelihood we therefore average the velocity distribution in v_{nb} over the distribution of the binary component v' given in Equation (6):

$$L_b(v|\sigma_0, \bar{v}) = \int_{-\infty}^{\infty} \frac{e^{-(v-v'-\bar{v})^2/2\sigma_0^2}}{\sqrt{2\pi\sigma_0^2}} \frac{f_b(\log|v'|)}{2|v'|\ln 10} dv'. \quad (8)$$

The factor of 2 in the denominator arises from the fact that $f_b(\log|v|)$ is normalized in $\log|v|$, whereas the likelihood is normalized in v (allowing for positive and negative velocities).

By taking the second moment of the velocity distribution in Equation (7), one obtains the result that

$$\sigma^2 = \sigma_0^2 + B\sigma_b^2, \quad (9)$$

where σ is the measured dispersion and σ_b is the binary dispersion found by taking the second moment of the binary velocity distribution $f_b(v) = f_b(\log|v|)/2|v|\ln 10$. As in the usual case, it can be shown that given a normally distributed measurement error with dispersion σ_m, one only needs to make the replacement $\sigma_0^2 \longrightarrow \sigma_0^2 + \sigma_m^2$ in the above formulae.

Next, it is desirable to have a likelihood function for velocities measured at two different epochs. Since velocity changes of order km s^{-1} over a timescale of years is entirely negligible for nonbinary stars, the most fruitful approach is to use a likelihood in the difference Δv between the two velocities. Keeping in mind the log-normal behavior of the velocity distribution, we write the binary part of the likelihood as $g_b(\log|\Delta v|; \Delta t)$. As with the single-epoch velocity distribution $f_b(\log|v|)$, we use a Monte Carlo simulation to calculate $g_b(\log|\Delta v|; \Delta t)$. For each binary in the simulation we find Δv by evolving the orbital phase to its value after a time Δt and calculating the resulting change in velocity. In the absence of measurement error, the nonbinaries will have zero change in velocity, so the total likelihood can be written as

$$L(\Delta v|\Delta t, B) = (1 - B)\delta(\Delta v) + B\frac{g_b(\log|\Delta v|; \Delta t)}{2|\Delta v|\ln 10}. \quad (10)$$

If there is a normally distributed measurement error, the likelihood must be averaged over two Gaussians of widths $\sigma_{m,1}$ and $\sigma_{m,2}$ for the first and second velocity errors, respectively. A little calculation shows this to be equivalent to averaging over a single Gaussian with dispersion σ_{2e}, which is the equivalent two-epoch measurement error:

$$\sigma_{2e}^2 = \sigma_{m,1}^2 + \sigma_{m,2}^2, \quad (11)$$

$$L(\Delta v|\Delta t, B, \sigma_{2e}) = (1 - B)\frac{e^{-\Delta v^2/2\sigma_{2e}^2}}{\sqrt{2\pi\sigma_{2e}^2}} + B\int_{-\infty}^{\infty} \frac{e^{-(\Delta v-\Delta v')^2/2\sigma_{2e}^2}}{\sqrt{2\pi\sigma_{2e}^2}} \times \frac{g_b(\log|\Delta v'|; \Delta t)}{2|\Delta v'|\ln 10} d(\Delta v'). \quad (12)$$

Note that the likelihood is identical in form to that of Equations (7) and (8), since in both cases we are averaging the distribution over a Gaussian. Both the single- and multi-epoch likelihoods will be put to use in later sections.

3.2. Threshold Fraction of a Binary Population

A convenient observable quantity for characterizing a binary population is the *threshold fraction*, defined as the fraction F of stars in a sample which exhibit a change in radial velocity greater than a threshold Δv after a time Δt between measurements. For $\Delta v > 1$ km s^{-1}, this fraction is typically smaller than 0.2, so the threshold *number* (given by $n = NF$, where N is the number of stars) follows a Poisson distribution with mean $\bar{n} = N\bar{F}$. Therefore, the distribution of F is characterized by a single number, the mean threshold fraction $\bar{F}$, and the expected error can be estimated. In particular, the error in F is approximately $\sqrt{\bar{F}/N}$ (Appendix B). For notational simplicity, for the remainder of this paper we will refer to the mean threshold fraction $\bar{F}$ as simply the threshold fraction F (without the bar), with the understanding that the observed threshold fraction will have a Poisson scatter about this value.

Despite its straightforward definition, there are two difficulties in measuring the threshold fraction from actual data sets. First, often there does not exist a common time interval Δt between measurements in the sample, but rather several time intervals for various subsets of stars. Furthermore, different velocity measurements have their own associated measurement errors and this in turn affects the measured value of F. The latter issue can be dealt with in an approximate way by using the median (or other robust location estimator) of the measurement error of the sample, in terms of which $\bar{\sigma}_{2e} = \bar{\sigma}_m\sqrt{2}$

(Equation (11)). However, both problems can be surmounted more rigorously by estimating the error-free threshold fraction F_0 via a Bayesian or maximum-likelihood approach. By using the likelihood in Δv defined in Equation (12), the threshold fraction at a particular threshold and time interval can be estimated even if measurements were taken at various epochs—moreover, the inferred threshold fraction F_0 is free of measurement error. This method will be demonstrated in Section 4.

The threshold fraction without measurement error, which we denote by F_0, can be expressed in terms of the binary two-epoch velocity distribution $g_b(\log|\Delta v|;\Delta t)$ by taking the integral of Equation (10) with respect to $|\Delta v'|$ from a threshold Δv to ∞:

$$F_0(\Delta v|\Delta t, B) = B\int_{\Delta v}^{\infty}\frac{g_b(\log|\Delta v'|;\Delta t)}{|\Delta v'|\ln 10}d|\Delta v'|. \quad (13)$$

Note that in the absence of measurement error, the threshold fraction F_0 scales linearly with the binary fraction B. The threshold fraction with measurement error is likewise obtained by taking the integral of Equation (12) from Δv to ∞, with the result

$$F(\Delta v|\Delta t, B, \sigma_{2e}) = (1-B)\mathrm{erfc}\left[\frac{\Delta v}{\sqrt{2}\sigma_{2e}}\right] + B\int_{-\infty}^{\infty}\mathrm{erfc}\left[\frac{\Delta v-\Delta v'}{\sqrt{2}\sigma_{2e}}\right] \times\frac{g_b(\log|\Delta v'|;\Delta t)}{2|\Delta v'|\ln 10}d(\Delta v'), \quad (14)$$

where σ_{2e} is the two-epoch measurement error given by Equation (11). Note that in the limit as $\sigma_{2e}\to 0$, the first term goes to zero while the complementary error function in the integrand reduces to a step function $2\Theta(\Delta v'-\Delta v)$, so that Equation (14) reduces to Equation (13) as expected.

4. CONSTRAINING PROPERTIES OF A BINARY POPULATION BY MULTI-EPOCH OBSERVATIONS

In this section, we investigate how properties of a population of binary stars affect the observed velocity distribution measured at two or more epochs. Specifically, we explore how our proposed observable, the threshold fraction F (Section 3.2), will be affected by changes in the underlying binary fraction B, absolute magnitude limit, stellar age, size of the measurement error, and time interval between measurements. We will also demonstrate how the binary fraction B can be inferred by a likelihood analysis and show how this leads to a better determination of the threshold fraction F. We first simulate binary populations that follow our fiducial period distribution (observed in solar neighborhood binaries) and then explore how the inferred binary fraction is affected if the assumed period distribution parameters are incorrect. Unfortunately, and as we discuss more fully in the next section, the effect of changing the binary fraction B on the observed binary velocities can be mimicked closely by altering the assumed distributions of orbital parameters (e.g., by changing the period distribution parameters $\mu_{\log P}, \sigma_{\log P}$). While this is bad news for any attempt at constraining the underlying properties of a galaxy's binary population in full generality, it turns out to be good news for correcting the observed velocity dispersion for the effects of binary orbital motion, as we will show in Section 5.

For the moment, let us first make the rather optimistic assumption that the distribution of binary orbital periods is approximately universal, so that it follows our fiducial choice $\mu_{\log P} = 2.24$, $\sigma_{\log P} = 2.3$ (Section 2). Before launching into the full-fledged calculation, one would like to estimate how well the fiducial binary fraction B can be constrained for a given sample, or conversely, how many stars are required to constrain B by a certain amount. To simplify matters, let us assume we have a data set where the two epochs have the same time interval Δt between them and the same measurement error σ_m. The equivalent two-epoch measurement error is then $\sigma_{2e} = \sigma_m\sqrt{2}$ (Equation (11)).

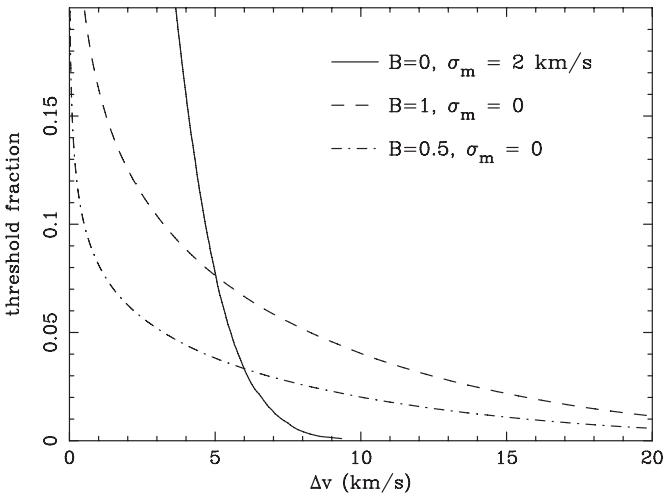


Figure 2. Threshold fraction $F(\Delta v|\Delta t, B, \sigma_m)$ defined as the fraction of stars with observed change in velocity greater than a threshold Δv after a time interval Δt between measurements. The solid curve has no binaries ($B = 0$) and a measurement error $\sigma_m = 2$ km s^{-1}. The other curves are plotted from a Monte Carlo simulation for binary fractions $B = 1$ and 0.5, with $\Delta t = 1$ yr and no measurement error. The binary population here follows the fiducial (solar neighborhood) period distribution. The stellar population has an age $t_g = 10$ Gyr and the absolute magnitude limit $M_{\rm lim} = 1$. Given a measurement error σ_m, the binary fraction can be constrained for thresholds $\Delta v \gtrsim \Delta v_{\rm tail}$, where $\Delta v_{\rm tail}$ is the point of intersection where $F(\Delta v|\Delta t, B = 0, \sigma_m) = F(\Delta v|\Delta t, B, \sigma_m = 0)$. For a given binary fraction B, the total threshold fraction without measurement error is given by $B \times F(\Delta v|\Delta t, B = 1, \sigma_m = 0)$.

In Figure 2, we plot the threshold fraction of a binary population as outlined above but with zero measurement error, denoted by $F_0(\Delta v|\Delta t, B)$ (Equation (13)), for two different binary fractions $B = 0.5$ and $B = 1$. These graphs are produced by a Monte Carlo simulation with an absolute magnitude limit $M_{\rm lim} = 1$ and stellar age $t_g = 10$ Gyr. We also plot the threshold fraction from measurement error *only*, i.e., without binaries and with $\sigma_m = 2$ km s^{-1}; this is given by the complementary error function (Equation (14) with $B = 0$). Near the point of intersection $\Delta v_{\rm tail}$ where $F_0(\Delta v_{\rm tail}|\Delta t, B) = F(\Delta v_{\rm tail}|\Delta t, B = 0, \sigma_{2e})$, the effect of binary stars becomes noticeable over the measurement error. Since the Poisson errors are larger at higher velocity thresholds, to first approximation we can say that the error-free threshold fraction F_0 is best constrained at thresholds near $\Delta v_{\rm tail}$. It follows that the fiducial binary fraction will be constrained by the stars with $\Delta v \gtrsim \Delta v_{\rm tail}$. (For a rough approximation, one can also use $\Delta v_{\rm tail} \approx 2\sigma_{2e} \approx 2\sqrt{2}\sigma_m$.) A little algebra (see Appendix B) shows that to constrain the binary fraction within a fractional accuracy of ϵ_b, the number of stars required is approximately

$$N(\epsilon_b) \approx \frac{1}{\bar{F}(\Delta v_{\rm tail})}\left(\frac{2B}{\epsilon_b}\right)^2. \quad (15)$$

In Figure 3, we graph the approximation formula for different

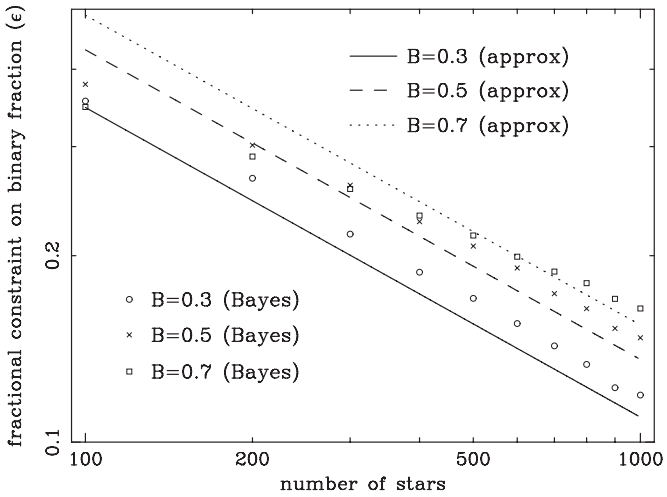


Figure 3. Fractional constraint ϵ_b on the binary fraction B, defined by Equation (15). In this plot, we show a measurement error $\sigma_m = 2.0$ km s^{-1}. For comparison we also plot the 95% confidence interval obtained by a Bayesian analysis of simulated two-epoch data from a random sample of N stars, averaged over a hundred realizations. A uniform prior is assumed for B.

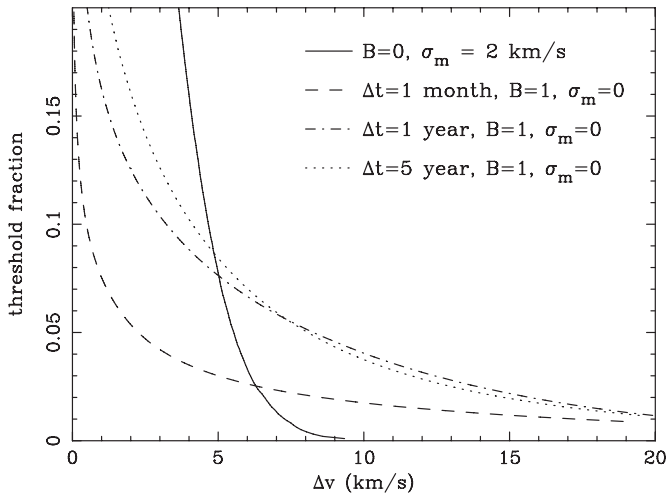


Figure 4. Threshold fraction $F(\Delta v|\Delta t, B, \sigma_m)$, defined as the fraction of stars with observed change in velocity greater than a threshold Δv after a time interval Δt between measurements. The solid curve has no binaries ($B = 0$) and a measurement error $\sigma_m = 2$ km s^{-1}. The other curves are plotted from a Monte Carlo simulation for different time intervals Δt, with no measurement error and a binary fraction $B = 1$. The binary population here follows the fiducial (solar neighborhood) period distribution. The stellar population has an age $t_g = 10$ Gyr and the absolute magnitude limit $M_{\rm lim} = 1$. Given a measurement error σ_m, the binary fraction can be constrained for thresholds $\Delta v \gtrsim \Delta v_{\rm tail}$, where $\Delta v_{\rm tail}$ is the point of intersection where $F(\Delta v|\Delta t, B = 0, \sigma_m) = F(\Delta v|\Delta t, B, \sigma_m = 0)$. For a given binary fraction B, the total threshold fraction without measurement error is given by $B \times F(\Delta v|\Delta t, B = 1, \sigma_m = 0)$.

values of B and compare to the 95% confidence interval in the binary fraction inferred by a Bayesian analysis of the simulated data (described later in this section). As is evident for the $B = 0.7$ curve, the approximation formula differs for high binary fractions because $B > 1$ is not allowed in the Bayesian analysis. The approximation formula is discussed further in Appendix B.

It is natural to ask whether better constraints can be obtained by lengthening the time interval between measurements. To address this, in Figure 4 we plot the threshold fraction $\bar{F}(\Delta v|\Delta t, B = 1, \sigma_{2e} = 0)$ produced by a Monte Carlo simulation with an absolute magnitude limit $M_{\rm lim} = 1$ and stellar age $t_g = 10$ Gyr for different time intervals Δt. We find that for a measurement error $\sigma_m = 2$ km s^{-1}, the observable threshold fraction steadily increases as Δt is increased, until roughly $\Delta t = 1$ yr. This result depends somewhat on the mass ratio distribution, since higher mass ratios result in higher velocities for given orbital periods of order Δt. If the mass ratio distribution in Duquennoy & Mayor (1991) is assumed for all periods (as opposed to the uniform distribution we adopt for $P < 3000$ days), the observable threshold fraction increases until roughly $\Delta t = 2$ yr. In any case, unless the measurement error is smaller than 2 km s^{-1}, little is gained by extending the interval beyond 1–2 yr.

In Figure 5, we plot the threshold fraction for different absolute magnitude limits $M_{\rm lim} = 0, 3$ and stellar ages $t_g = 1$ Gyr and 10 Gyr. Extending the magnitude limit to fainter magnitudes increases the threshold fraction because there is a greater contribution from smaller stars with less binary suppression due to Roche-lobe overflow. The threshold fraction is also higher for a younger stellar population because of their larger mass at a given stage of stellar evolution, which produces higher orbital velocities. We find that the threshold fraction changes little for ages between 2 and 10 Gyr; as the stellar age t_g is reduced from 2 Gyr, however, the threshold fraction steadily increases.

To estimate the binary fraction in a two-epoch sample, we use the likelihood function $L(\Delta v|\Delta t, B, \sigma_{2e})$ (Equation (12)). For the sake of illustration, we analyze a simulated galaxy with binary fraction $B = 0.5$ but with a different period distribution from that of the solar neighborhood. We choose the period

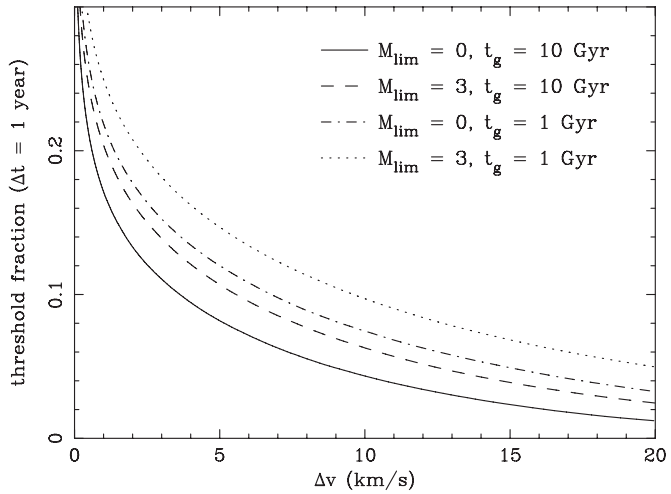


Figure 5. Threshold fraction $F(\Delta v|\Delta t, B, \sigma_m)$, defined as the fraction of stars with observed change in velocity greater than a threshold Δv after a time interval Δt year between measurements. Here we have picked $\Delta t = 1$ yr, binary fraction $B = 1$, and no measurement error ($\sigma_m = 0$). We plot the threshold fraction for different absolute magnitude limits $M_{\rm lim} = 0$, 3 and stellar ages $t_g = 1$ Gyr, 10 Gyr. The binary population here follows the fiducial (solar neighborhood) period distribution. For a given binary fraction B, the total threshold fraction is given by $F(\Delta v|\Delta t, B, \sigma_m = 0) = B \times F(\Delta v|\Delta t, B = 1, \sigma_m = 0)$.

distribution parameters $\mu_{\log P} = 1.5$, $\sigma_{\log P} = 2$ (P in years) for this galaxy. The simulated data sample consists of 300 stars, each with two velocity measurements taken $\Delta t_{\rm data} = 2$ yr apart, and a measurement error of 2 km s^{-1}. First, we assume the fiducial model (with $\mu_{\log P} = 2.24$, $\sigma_{\log P} = 2.3$, which is incorrect for this galaxy) and, assuming a uniform prior in the binary fraction B, generate a posterior in the binary fraction. We then repeat this procedure using the correct period distribution parameters $\mu_{\log P}$, $\sigma_{\log P}$ in our model, whose values are given above. The resulting posteriors are plotted in Figure 6. This figure shows that

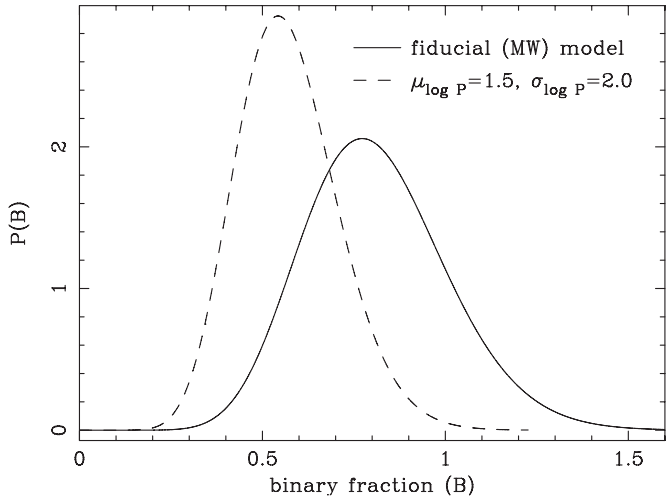


Figure 6. Posterior probability distribution of the binary fraction B of a simulated galaxy with binary fraction $B = 0.5$ and with a period distribution characterized by $\mu_{\log P} = 1.5, \sigma_{\log P} = 2$ (P in years). The simulated data sample consisted of 300 stars, each with two velocity measurements taken $\Delta t_{\rm data} = 2$ yr apart. The solid curve is the posterior calculated assuming the fiducial (solar neighborhood) model, which is incorrect for this galaxy. The dashed curve uses the correct model, with the same period distribution parameters $\mu_{\log P}, \sigma_{\log P}$ given above.

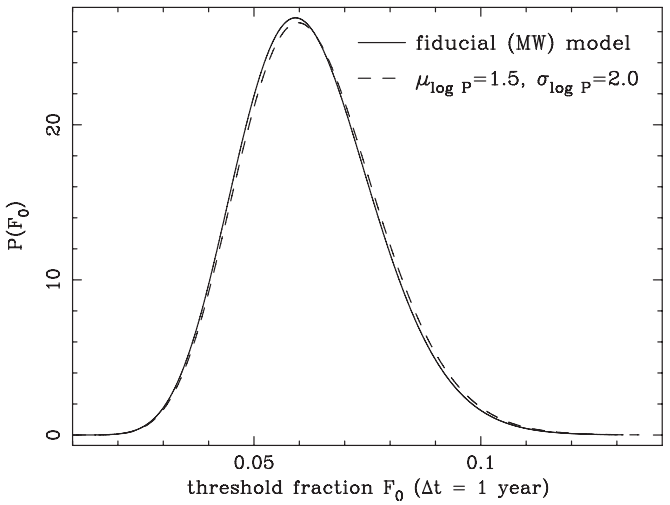


Figure 7. Posterior probability distribution of the threshold fraction F_0 of a simulated galaxy with binary fraction $B = 0.5$ and with a period distribution characterized by $\mu_{\log P} = 1.5$, $\sigma_{\log P} = 2$ (P in years). The threshold fraction $F_0(\Delta v|\Delta t)$ is defined as the fraction of stars with change in velocity greater than a threshold Δv after a time interval Δt year between measurements. Here we chose a threshold $\Delta v = 5$ km s^{-1} and time interval $\Delta t = 1$ yr; the correct threshold fraction for this galaxy is $F_0 \approx 0.05$. The simulated data sample consisted of 300 stars, each with two velocity measurements taken $\Delta t_{\rm data} = 2$ yr apart, and a measurement error of 2 km s^{-1}. The solid curve is the posterior calculated assuming the fiducial (solar neighborhood) model, which is incorrect for this galaxy. The dashed curve uses the correct model, with the same period distribution parameters $\mu_{\log P}, \sigma_{\log P}$ given above. Note that the correct threshold fraction can be recovered even if the wrong model is assumed (in this case, the fiducial model).

the binary fraction B is a highly model-dependent quantity, and given the unknown nature of the period distribution of binaries outside the solar neighborhood, the inferred binary fraction must be taken with a grain of salt. However, the fiducial binary fraction can still be used as a relative indicator of the fraction of observable binaries, as long as it is interpreted in reference to the fraction observed in a binary population following the fiducial (solar neighborhood) distributions of orbital parameters.

Although the inferred binary fraction B is very model dependent, this analysis is still useful in that it leads to a better determination of the threshold fraction F_0, which is more directly observable than the binary fraction. To see this, we use the Monte Carlo to generate the binary threshold fraction $F_b \equiv F_0(\Delta v|\Delta t, B = 1)$ of each model, for a threshold $\Delta v = 5$ km s^{-1}, time interval $\Delta t = 1$ yr, and zero measurement error. We then transform each posterior in Figure 6 from B to the threshold fraction F_0 according to $F_0 = BF_b$ (Equation (13)). The renormalized posteriors $P(F_0)$ are plotted in Figure 7; the correct threshold fraction for this galaxy is $F_0 \approx 0.05$. Note that the correct threshold fraction can be recovered even if the wrong model is assumed (in this case, the fiducial model).

Since each star is weighted appropriately by its measurement error and even the stars with velocities smaller than the threshold Δv are used in the likelihood analysis, the error in the threshold fraction is smaller than if F were measured directly, especially for higher velocity thresholds. The approximate error in the threshold fraction F_0 estimated by this technique is derived in Appendix B and given by Equation (B3). Furthermore, the threshold fraction at $\Delta t = 1$ yr is recovered even though the data were taken with a time interval of $\Delta t_{\rm data} = 2$ yr. More generally, the threshold fraction for a specific time interval can be recovered by the likelihood analysis even if the data are taken at various different epochs with different measurement errors. As we will show in Section 5, the threshold fraction can be used to correct the measured velocity dispersion of a sample for the effect of binary motion.

Finally, one may naturally wonder: by how much are the constraints improved by including more than two epochs in the analysis? To address this question, we do a similar calculation on simulated three-epoch data using the Monte Carlo to generate a three-epoch likelihood, $L(\Delta v_{21}, \Delta v_{32}|\Delta t_{21}, \Delta t_{32}, B, \sigma_{3e})$, where the indices refer to three velocity measurements v_1, v_2, and v_3, and σ_{3e} is the three-epoch measurement error defined analogously to Equation (11). We generate a data sample of 500 stars from a simulated galaxy with binary fraction $B = 0.5$ and the period distribution parameters taking their fiducial values. For comparison, we generate a posterior $P(B)$ from the two-epoch calculation which ignores v_3. The results are plotted in Figure 8. The velocity measurements v_1 and v_2 were taken one year apart, while v_2 and v_3 were taken 10 yr apart. While the most probable inferred binary fraction $B_{\rm fit}$ did improve in this case, the 95% confidence limits are only improved by ≈25%. The fractional improvement in the confidence limits is even less for smaller data sets; this is because in a sample of a few hundred stars, there is significant scatter in the binary fraction, and the inferred binary fraction $B_{\rm fit}$ in fact has a significant probability of becoming worse when more epochs are added to the sample. To constrain the binary and/or threshold fraction, we therefore find it a more profitable strategy to make two-epoch measurements over a larger sample of stars, as opposed to adding more repeat measurements over an existing sample (assuming a similar number of overall measurements in either case).

A possible complicating factor in the above analysis is that selection criteria for making repeat measurements can bias the inferred threshold fraction and binary fraction. If stars whose spectra yield multiple peaks in the cross-correlation function are singled out for repeat measurements, the multi-epoch sample may have an inordinately high binary fraction compared to the overall stellar population. This selection bias is probably not significant in red giant stars due to the typically large difference

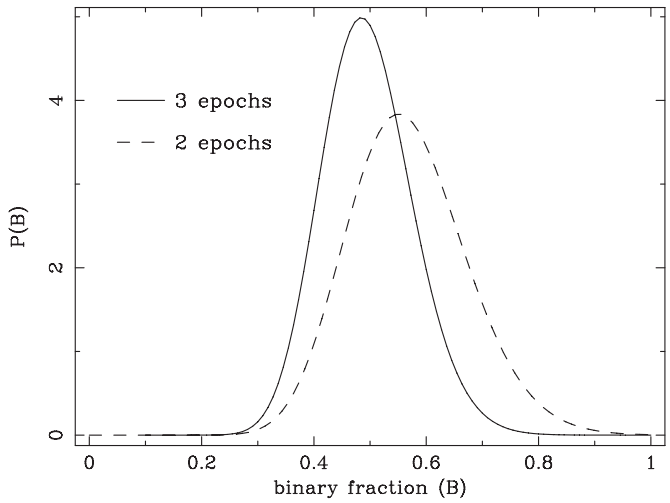


Figure 8. Posterior probability distribution for binary fraction of a simulated galaxy with 500 stars and binary fraction $B = 0.5$. The solid curve is calculated from three velocity measurements, whereas the dotted curve uses only the first two velocity measurements. v_1 and v_2 were taken one year apart, while v_2 and v_3 were taken 10 yr apart.

in luminosity between primary and secondary star. However, in samples that contain a significant fraction of main-sequence stars the bias may be more problematic, although an upper bound on the binary dispersion can still be obtained.

5. CORRECTING THE OBSERVED VELOCITY DISPERSION FROM MULTI-EPOCH DATA

5.1. Degeneracy of Binary Fraction with Period Distribution Parameters

In the previous section, we demonstrated how uncertainties in the underlying period distribution of a binary population can undermine our ability to constrain the binary fraction from multi-epoch data. Here we demonstrate that if our goal is to correct the observed velocity dispersion for the effects of binary stars, the degeneracy between the period distribution parameters and binary fraction is quite useful: regardless of the precise nature of the binary population, its effect on the observable threshold fraction F can be directly related to the associated correction in the observed velocity dispersion in a model-independent way.

The important degeneracy arises from the log-normal behavior of the binary orbital velocity distribution $f_b(\log|v|)$ (Equation (6)), which we discussed at the end of Section 2. Binary orbital velocities of order km s^{-1} make the largest contribution to the velocity dispersions of dwarf galaxies. For these velocities, the value of $\log|v|$ is far from the mean of the log-normal, which is approximately -1 for a magnitude limit $M_{\rm lim} = 1$ and age $t_g = 10$ Gyr. The exponent of the log-normal is approximately linear over the scale of km s^{-1}, so we can write it as $-\beta - \alpha \ln|v|$. Therefore, the binary part of the velocity distribution can be written as $f(v) \propto Be^{-\beta}|v|^{-1-\alpha}$, where B is the binary fraction. If the mean binary period $\mu_{\log P}$ is varied, the log-normal is offset in log space so to good approximation only β changes; therefore, the velocity distribution $f(v)$ changes by a constant factor over the scale of km s^{-1}. If the dispersion of the period distribution $\sigma_{\log P}$ is varied, both the offset β and the slope α change; however, the slope changes by a relatively small amount for $\sigma_{\log P}$ ranging from 1 to 3 (its viable range of values—see Section 7), so again the velocity distribution changes by an approximately constant factor. The essential point is that if the parameters $\sigma_{\log P}$ and $\mu_{\log P}$ are varied, they change the velocity distribution by an amount which is nearly the same over the scale of several km s^{-1}—in other words, they behave similarly as if the binary fraction were changed. This is also true of the magnitude limit and stellar age, which effectively change the mean of the log-normal and therefore behave similarly to $\mu_{\log P}$. We therefore conclude that the parameters $\mu_{\log P}$, $\sigma_{\log P}$, magnitude limit $M_{\rm lim}$, and stellar age t_g are all nearly degenerate with binary fraction over the scale of km s^{-1}.

The degeneracy of the period distribution parameters with binary fraction also holds for the two-epoch velocity distribution $g_b(\log|\Delta v|; \Delta t)$, since this also has a log-normal form for km s^{-1} velocities. By Equation (14), therefore, the same degeneracy holds for the threshold fraction F. The effect of this degeneracy on the threshold fraction and its implications for constraining the binary distribution of periods will be explored in further detail in Section 7.

5.2. Correlation of Binary Velocity Dispersion with Observed Threshold Fraction

In this section, we will consider the threshold fraction F_0 with a fixed time interval of 1 yr and without measurement error, i.e., $F_0 = F(\Delta v|\Delta t = 1\,{\rm yr}, \sigma_{2e} = 0)$ (Section 3.2). There is no loss of generality in this; as we demonstrated in Section 4, the threshold fraction for any given time interval Δt can be estimated by a likelihood analysis even if measurements are taken at various different epochs and with various different measurement errors. However, if the threshold fraction F is measured directly for a fixed time interval, it is necessary to account for the effect of measurement error on F; we will address this later in the section.

In the absence of measurement error, by definition F_0 scales linearly with the binary fraction B (Equation (13)). Furthermore, because of the near-degeneracy of the period distribution parameters with binary fraction, F_0 also scales linearly with $\mu_{\log P}$ and $\sigma_{\log P}$ to good approximation over their viable range of values (roughly 1–3 with P in years; see Section 7). The essential point is that a similar relationship holds for the velocity dispersion if a high-velocity cutoff is used, e.g., at $v_c = 3\sigma$, since the degeneracy approximately holds for velocities $v < v_c$. It follows that if velocity outlier stars are excluded in determining velocity dispersion, the extra dispersion due to binaries can be determined from the threshold fraction F_0 with reasonable confidence even if the parameters B, $\mu_{\log P}$, and $\sigma_{\log P}$ are entirely unknown.

We demonstrate this by simulating galaxies with various intrinsic dispersions and characterized by different binary populations. The measured dispersion σ is calculated by iteratively discarding stars with velocities larger than 3σ; on the first iteration, the biweight is used to estimate the dispersion, and the dispersion is then calculated on every subsequent iteration until all the remaining stars have velocities that lie within 3σ. In order to make the statistical error small, we used a very large "sample" of 100,000 stars. We will also calculate the threshold fraction F_0 for the same data set, for which we choose a threshold $\Delta v = 5$ km s^{-1} and time interval $\Delta t = 1$ yr.

To demonstrate the effect, we first simulate a galaxy with an intrinsic dispersion of 4 km s^{-1} and a binary population that follows the fiducial period distribution ($\mu_{\log P}$ and $\sigma_{\log P}$). In Figure 9, we plot the ratio σ/σ_0 of measured dispersion over the intrinsic dispersion as a function of binary fraction for a galaxy with an intrinsic dispersion of 4 km s^{-1}. The solid and dashed curves assume an absolute v-band magnitude limit $M_{\rm lim} = 1$ and $M_{\rm lim} = 3$, respectively. For the dot-dashed curve

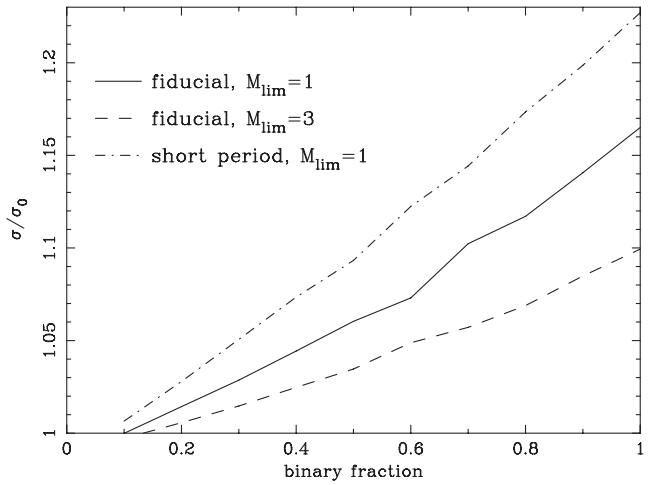


Figure 9. Ratio of measured velocity dispersion σ over the intrinsic dispersion σ_0, plotted with respect to binary fraction in a simulated galaxy of intrinsic dispersion $\sigma_0 = 4$ km s^{-1}. For the solid and dashed curves, the binary population follows the fiducial period distribution ($\mu_{\log P} = 2.24$, $\sigma_{\log P} = 2.3$, P in years) while different absolute magnitude limits are assumed. For the dot-dashed curve, the binary population has a mean period of 10 yr and a more peaked period distribution ($\mu_{\log P} = 1$, $\sigma_{\log P} = 0.5$). The measured dispersions were calculated by an interactive 3σ-clipping routine.

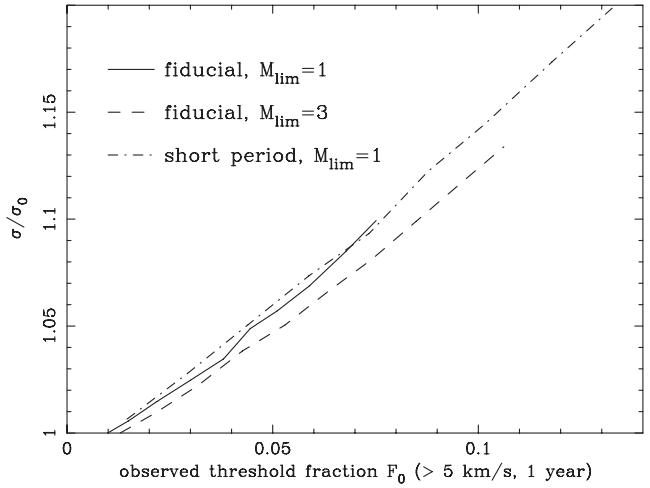


Figure 10. Same as Figure 9, but plotted with respect to threshold fraction F_0. The threshold fraction F_0 denotes the fraction of stars with observed change in velocity greater than a threshold $\Delta v = 5$ km s^{-1} after a time $\Delta t = 1$ yr between measurements, assuming zero measurement error.

we have chosen a binary population with a shorter mean period of 10 yr and a more peaked period distribution ($\mu_{\log P} = 1$, $\sigma_{\log P} = 0.5$). Note that for a given binary fraction, the effect on the measured dispersion depends on the period distribution as well as the magnitude limit. However, if we plot the same points with respect to the threshold fraction F_0, as in Figure 10, we find that the observed threshold fraction can be mapped in a nearly one-to-one way to the intrinsic dispersion regardless of the nature of the period distribution or magnitude limit.

Next we repeat the procedure over a grid of values for the parameters B, $\mu_{\log P}$, and $\sigma_{\log P}$, and for each point we plot the ratio σ/σ_0 with respect to the threshold fraction F_0. The results are plotted in Figure 11, again for galaxies with intrinsic dispersions of 4, 7, and 10 km s^{-1}. We see that for each group, the graph shows a tight correlation for all but the most extreme values of the period distribution parameters. In plotting these points we varied B from 0.2 to 1, $\mu_{\log P}$ from -1 to 4, and $\sigma_{\log P}$ from 0.5 to 4 (with P in years). The lowermost points of each

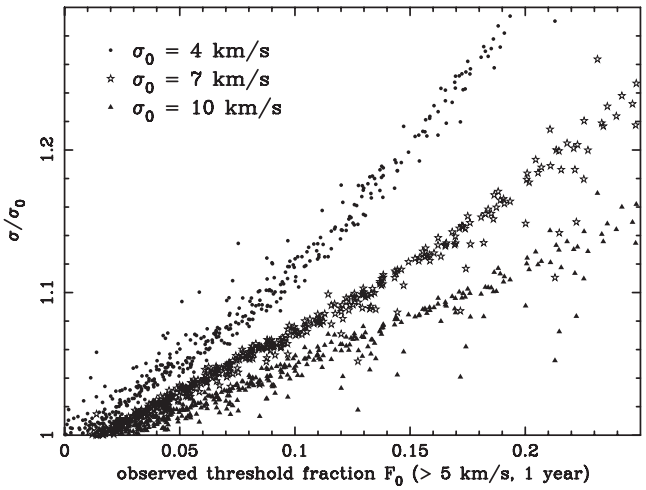


Figure 11. Ratio of measured velocity dispersion σ over the intrinsic dispersion σ_0, plotted with respect to threshold fraction F_0 for different binary populations in galaxies of intrinsic dispersions $\sigma_0 = 5$, 7, and 10 km s^{-1}. The measured dispersions were calculated by an interactive 3σ-clipping routine, and the threshold fraction F_0 denotes the fraction of stars with observed change in velocity greater than a threshold $\Delta v = 5$ km s^{-1} after a time $\Delta t = 1$ yr between measurements, assuming zero measurement error. Each point represents a different binary population with its own binary fraction and period distribution; we plotted the points over a grid of values, with binary fraction B ranging from 0.2 to 1, $\mu_{\log P}$ from -1 to 4 (in $\log(P/\mathrm{yr})$), and $\sigma_{\log P}$ from 0.5 to 4. We show at the end of Section 5 how F_0 can be inferred from observations accounting for measurement errors.

group are the points for which $\sigma_{\log P}$ has its smallest value of 0.5, producing only a relatively small number of short-period binaries. The uppermost points are the points for which $\mu_{\log P}$ has its smallest value, so the period distribution is shifted toward short periods. For these extreme values, the velocity distribution becomes distorted into a distinctly non-Gaussian shape so these can be considered improbable configurations.

Although we have singled out the period distribution parameters to demonstrate the correlation in Figure 11, the same degeneracy with binary fraction exists for the distributions of *all* the binary orbital parameters. For example, by varying the ellipticity distribution parameters $\bar{e}$, σ_e (Section 2) in addition to B, $\mu_{\log P}$, and $\sigma_{\log P}$, we find that the tight correlation in Figure 11 is unchanged; although the amount of scatter increases slightly, the correction still holds to within a few percent accuracy. The same considerations apply when parameterizing the distribution of mass ratios. To within a few percent accuracy, the correlation between the observed threshold fraction and the dispersion introduced by binaries holds independent of the nature of the underlying binary population.

The correlation in Figure 11 also remains the same regardless of the magnitude limit and the age of the stellar population, again because of the degeneracy with binary fraction discussed above. However, as the absolute magnitude limit is extended to $M_{\mathrm{lim}} > 3$, the degeneracy becomes weaker so that smaller σ/σ_0 values are allowed for a given threshold fraction. Therefore in samples containing main-sequence stars, the dispersion cannot be corrected with as great an accuracy, although an upper bound to the correction can still be obtained.

5.3. Fitting Functions to Correct Velocity Dispersion for Binary Motion

The points plotted in Figure 11 are well fit by a superposition of a line and an exponential function, provided that outlier points are discarded. To define "outliers," first we divide the

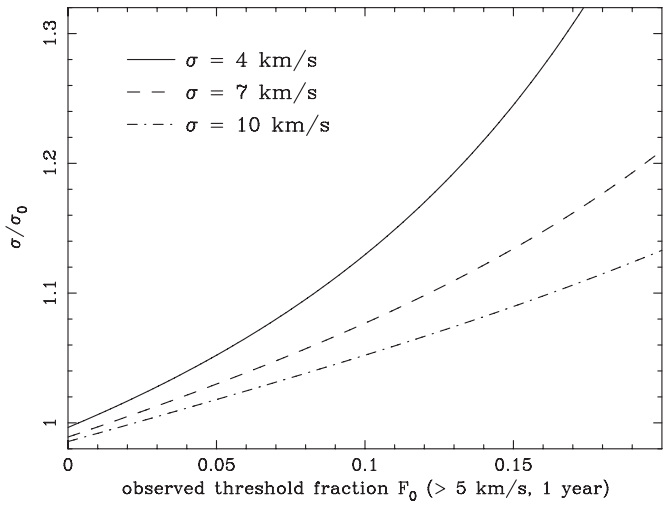


Figure 12. Ratio of measured velocity dispersion σ over the intrinsic dispersion σ_0, plotted with respect to threshold fraction F_0 for *fixed measured dispersions* σ. These curves were found by fitting graphs like those shown in Figure 11, then transforming to fixed values of the measured dispersion σ. The measured dispersions in Figure 11 are calculated by an interactive 3σ-clipping routine, and the threshold fraction F_0 denotes the fraction of stars with change in velocity greater than a threshold $\Delta v = 5$ km s^{-1} after a time $\Delta t = 1$ yr between measurements, with zero measurement error. We show at the end of Section 5 how F_0 can be inferred from observations accounting for measurement errors.

domain $F_0 \in [0, 0.2]$ into sections small enough so the graph is approximately linear within each section. We then further divide each section into two subsections and calculate the median and median absolute deviation (MAD) of the y-values of the points in each subsection. Next we draw lines through the two points defined by the median $\pm$ twice the MAD of each of the two subsections, taking the center of the subsection as their x-value. The plotted points that lie outside the region defined by these lines represent extreme and highly improbable period distributions and are therefore discarded. We fit the remaining points and repeat the procedure for galaxies of dispersions ranging from 3 to 12 km s^{-1}.

The plots in Figure 11 are not directly applicable to real data because each graph was plotted for a fixed intrinsic dispersion σ_0, which is unknown (and is in fact what we are attempting to calculate!). We therefore use our fits together with a root-finding procedure, interpolating the fitting parameters in σ_0, to draw similar graphs at fixed values of σ. A few resulting curves are plotted in Figure 12. Again, we find that these curves are well fit by a line plus exponential,

$$\frac{\sigma}{\sigma_0} = a(\sigma) + b(\sigma)F_0 + c(\sigma)[e^{F_0/0.1} - 1], \tag{16}$$

where F_0 is the threshold fraction at 5 km s^{-1}. We also find fitting functions for the parameters $a(\sigma)$, $b(\sigma)$, and $c(\sigma)$ which fit well for dispersions σ ranging from 4 km s^{-1} to 10 km s^{-1}. Defining $\Delta\sigma = \sigma - 4$ km s^{-1}, we find

$$a(\sigma) = 0.988e^{-0.0007\Delta\sigma}, \tag{17}$$

$$b(\sigma) = 0.576 - 0.08\Delta\sigma + 0.772(1 - e^{-0.1\Delta\sigma}), \tag{18}$$

$$c(\sigma) = 0.043e^{-0.247\Delta\sigma}. \tag{19}$$

These formulae hold for any magnitude limit and stellar age, and the threshold fraction F_0 refers here to a velocity threshold of 5 km s^{-1}, time interval of one year, and zero measurement error.

How are these formulae adjusted if a different velocity threshold is desired? Ideally, one should use the smallest possible threshold that is not significantly affected by measurement error—this will include the most stars and therefore have a smaller scatter compared to higher thresholds. To use a different threshold Δv, the value of F_0 in Equation (16) must be scaled by the ratio $F_0(\Delta v)/F_0(5$ km s$^{-1})$. This ratio can be calculated by using Monte Carlo realizations to plot the threshold fraction as a function of threshold, which we will do in Section 4 (Figures 2 and 4). Again, the degeneracy of magnitude limit and period distribution parameters with binary fraction ensures that this ratio is virtually independent of the model parameters and magnitude limit, provided one does not transform to thresholds that are too high (>10 km s^{-1}). We find that for thresholds in the range 4 km s^{-1} $< \Delta v <$ 10 km s^{-1}, the ratio $F_0(\Delta v)/F_0(5$ km s$^{-1})$ can be fit by the function

$$\frac{F_0(\Delta v)}{F_0(5\,\mathrm{km\ s^{-1}})} = a + be^{-\Delta v/\Delta v_s} \tag{20}$$

where the best-fit parameters are $a = 0.0725$, $b = 1.897$, and $\Delta v_s = 6.947$ km s^{-1}. Thus to find the correction to the dispersion in terms of a given velocity threshold Δv, one substitutes Equation (20) into Equation (16) so that the fit is in terms of $F_0(\Delta v)$.

5.4. Accounting for Measurement and Statistical Errors

Next, we address the issue of measurement error. As mentioned above, the best approach is to estimate the error-free threshold fraction F_0 by a Bayesian or maximum-likelihood approach (see Section 4), in which case measurement error need not be considered here. However, if F is calculated directly from the data, then measurement error must be considered in the above formulae. We find that given a two-epoch measurement error σ_{2e} (Equation (11)), the threshold fraction $F(\Delta v|\sigma_{2e})$ is related to the threshold fraction without measurement error $F_0(\Delta v)$ by a linear transformation—yet another consequence of the degeneracy between binary fraction and the period distribution parameters (see Appendix A for a derivation of this result). The transformation takes the following form:

$$F(\Delta v|\sigma_{2e}) \approx \mathrm{erfc}\left[\frac{\Delta v}{\sqrt{2}\sigma_{2e}}\right] + \beta\left(\frac{\Delta v}{\sigma_{2e}}\right) F_0(\Delta v). \tag{21}$$

This result (together with Equation (20)) has been tested by using the Monte Carlo simulation to produce plots similar to Figure 11 for different velocity thresholds and measurement errors. The approximate analytic form of β can also be derived (see Appendix A); using this together with the Monte Carlo plots to map β, we find that β is well fit by the following function:

$$\beta(x) = \left(1 + ae^{-\frac{x}{x_s}}\right)\left\{1 - \bar{\kappa}\cdot\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)\right\}, \tag{22}$$

where $x = \Delta v/\sigma_{2e}$ and the best-fit parameters are $a = 0.05$, $x_s = 5$, and $\bar{\kappa} = 1.3$. Thus to find the correction to the dispersion with a given measurement error, one substitutes the error-free threshold fraction F_0 in terms of $F(\Delta v|\sigma_{2e})$ (given by Equation (21)) into Equation (16). The effect of the measurement error on the dispersion must also be taken into account either by finding the maximum likelihood dispersion,

or approximately by making the substitution $\sigma^2 = \sigma^2_{\rm meas} - \bar{\sigma}^2_m$ into Equation (16), where $\sigma_{\rm meas}$ is the measured dispersion and $\bar{\sigma}_m$ is the mean measurement error of the sample. The latter approximation works well if the measurement error follows a distribution which is well-peaked about its mean value.

Finally, the statistical error in the intrinsic dispersion σ_0 determined by this procedure can be estimated by noting that for a two-epoch sample of N_{2e} stars, the Poisson error in the threshold fraction (corrected for measurement error) is $\delta F \approx \sqrt{F_0(\Delta v)/N_{2e}}$. Let us assume that $N_{2e} > 10$ and the dispersion is measured in a larger single-epoch sample of N stars; then we can make the approximation that the sampling errors in σ and F are weakly correlated so they add approximately in quadrature. Propagating the error using Equation (16) gives the sampling error in the intrinsic dispersion,

$$\left(\frac{\delta\sigma_0}{\sigma_0}\right)^2 \approx \frac{\left[\left(\frac{\sigma}{\sigma_0}\right)^2 + \left(\frac{\bar{\sigma}_m}{\sigma_0}\right)^2\right]^2}{2N} + \left|\frac{\sigma_0}{\sigma}\frac{\partial(\sigma/\sigma_0)}{\partial F_0}\right|^2 \frac{F_0(\Delta v)}{N_{2e}}, \quad (23)$$

where the second term in Equation (23) is the two-epoch sampling error. Note that we have assumed the distribution of measurement errors to be peaked about a mean value $\bar{\sigma}_m$, such that the measured dispersion is approximately given by $\sigma^2_{\rm meas} = \sigma^2 - \bar{\sigma}_m$. As an example, if the single-epoch sample contains $N = 500$ stars and the measured threshold fraction is $F_0 = 0.1$ in a two-epoch subset of $N_{2e} = 100$ stars, the intrinsic dispersion σ_0 can be determined to within $\approx$5% for dispersions $\sigma > 4$ km s^{-1}. Equation (23) was tested with Monte Carlo simulations and found to be accurate to within 1% for $N = N_{2e}$; in the above example, the formula is also accurate to within 1%. However, the fractional error in σ_0 cannot be made smaller than $\approx$1% due to the inherent uncertainty in the binary population model represented by the width of the plots in Figure 11. We find that for an observed threshold fraction $F = 0.1$, the two-epoch sampling error is smaller than the single-epoch error unless $N/N_{2e} > 5$. Given a measured threshold fraction $F(\Delta v)$, the two-epoch sampling error is larger relative to the single-epoch error for smaller measured dispersions.

5.5. Summary: Procedure for Correcting the Velocity Dispersion

To summarize, the velocity dispersion of a dwarf spheroidal sample can be corrected for binaries by the following method.

1. Measure the threshold fraction F for a particular threshold velocity and a time interval $\Delta t = 1$ yr. This can be done in two ways: the threshold fraction can be measured directly, in which case one should use the smallest possible threshold that is not unduly affected by measurement error; this is approximately $\Delta v \approx 2\bar{\sigma}_{2e}$, where $\bar{\sigma}_{2e} = \sqrt{2}\bar{\sigma}_m$ is the median two-epoch measurement error. An alternative (and more rigorous) approach is to estimate the measurement error-free threshold fraction F_0 by a likelihood analysis. This procedure is demonstrated in Section 4.
2. Measure the velocity dispersion $\sigma_{\rm meas}$ of the sample by an iterative 3σ-clipping routine. Correct the dispersion for measurement error to find the error-free dispersion σ.
3. If the chosen velocity threshold is different from 5 km s^{-1}, scale the threshold fraction F_0(5 km s^{-1}) in Equation (16) according to Equation (20).
4. If the threshold fraction is measured directly, one must use Equation (21) to relate the error-free threshold fraction $F_0(\Delta v)$ in Equation (16) to the threshold fraction with measurement error, $F(\Delta v|\sigma_{2e})$.
5. After substituting F_0(5 km s^{-1}) in terms of $F(\Delta v|\sigma_{2e})$ (given by steps 3 and 4 above) into Equation (16), use Equation (16) together with Equations (17)–(19) to find the intrinsic velocity dispersion σ_0. The statistical error in σ_0 can be estimated from Equation (23).

5.6. Upper Limit to the Expected Binary Dispersion in Dwarf Spheroidal Galaxies

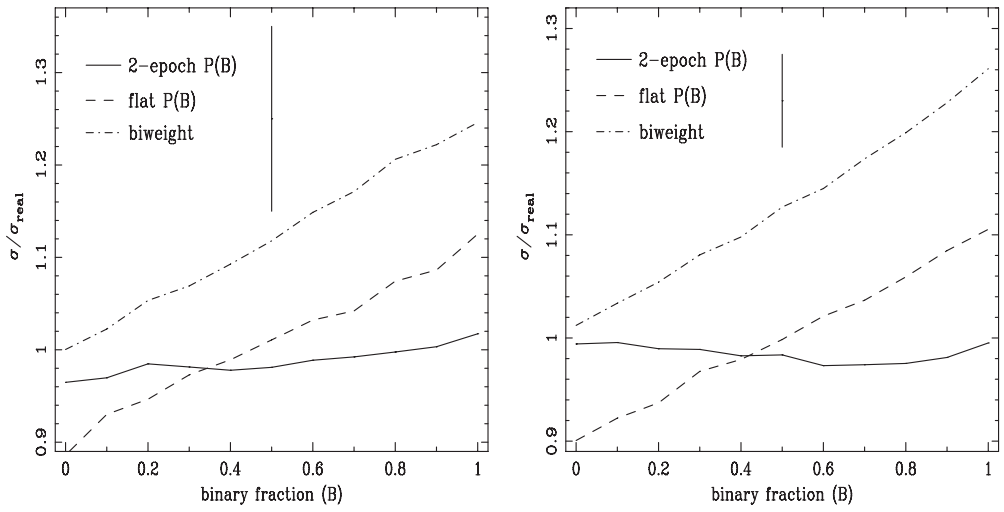
From Figure 11, we can estimate an upper limit to the possible dispersion introduced by binaries in dwarf spheroidal galaxies. First, we note that if a stellar sample consists entirely of red giants (i.e., magnitude limit $M_{\rm lim} \lesssim 1$) older than 1 Gyr, and if the distributions of binary orbital parameters are similar to those observed in the solar neighborhood, the threshold fraction F_0 of a sufficiently large sample cannot be larger than 0.1 even if the binary fraction is 1 (Figure 5). Even if the period distribution is shifted to shorter periods by two orders of magnitude, we still find $F_0 < 0.15$. By analyzing multi-epoch data in the Fornax, Carina, Sculptor, and Sextans dwarf spheroidals (Walker et al. 2009a) we find that they all have threshold fractions smaller than 0.12, while Fornax has F slightly larger than 0.1. We therefore conclude from Figure 11 that in dwarf spheroidal samples containing only red giant stars, the measured velocity dispersions are unlikely to be inflated by more than 20%.

In ultra-faint dwarf spheroidals, however, due to the small number of red giants it may become necessary to extend the sample to fainter magnitudes beyond the red giant branch (Geha et al. 2009). For absolute magnitude limits as high as $M_{\rm lim} = 6$, we find that the threshold fraction F_0 is unlikely to be greater than 0.2 unless its period distribution is shifted to significantly shorter periods (by an order of magnitude) than that of the solar neighborhood. We therefore conclude from Figure 11 that the measured velocity dispersions of dwarf spheroidals greater than 4 km s^{-1} are unlikely to be inflated by more than 30%. However, even if the dispersion of a particular galaxy is inflated by more than 30%, the correction due to binaries can be readily discerned by making repeat measurements and applying the method outlined above.

6. BAYESIAN ANALYSIS OF SINGLE-EPOCH VELOCITY DATA

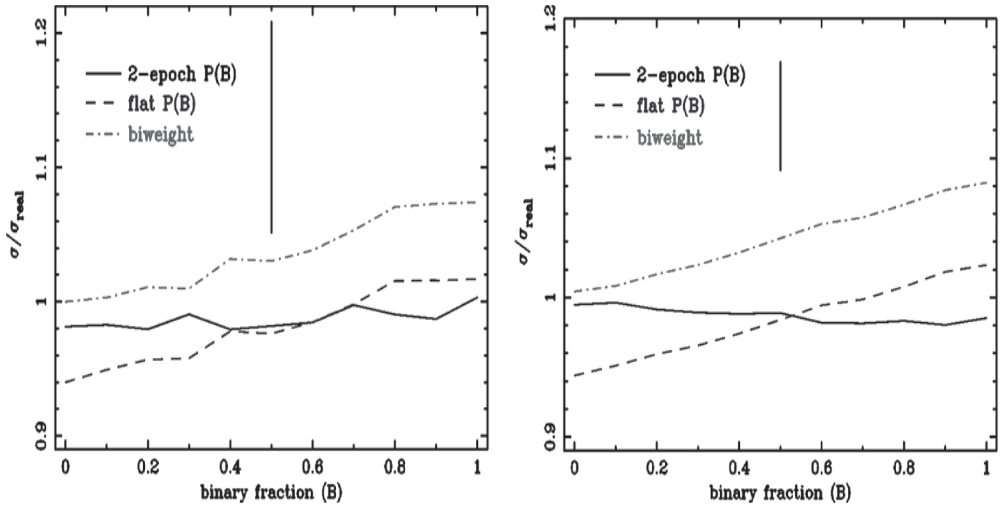
In this section, we discuss the problem of constraining properties of a binary population with radial velocity measurements taken at a single epoch. We approach this by fitting the likelihood for binary stars (Equation (7)) to a simulated data sample generated from a Monte Carlo simulation for galaxies with dispersions 4 km s^{-1} and 10 km s^{-1}. Both galaxies were given a binary fraction $B = 0.5$ and period distribution parameters equal to their fiducial values $\mu_{\log P} = 2.24$, $\sigma_{\log P} = 2.3$ (P in years). The simulated velocities were generated with a measurement error of 2 km s^{-1}. Using Equation (7) together with Bayes' theorem, marginal posterior distributions in the dispersion σ and binary fraction B can be obtained. The maxima of these distributions are taken as the best-fit values $\sigma_{\rm fit}$ and $B_{\rm fit}$.

In the absence of any constraint on the binary fraction from multi-epoch data, we assume a uniform prior in the binary fraction. However, if repeat measurements do exist for some subset of the data sample, these can be used to better constrain



(a) $\sigma_{real} = 4$ km/s, 100 stars (b) $\sigma_{real} = 4$ km/s, 500 stars

Figure 13. Best-fit values for the dispersion $\sigma_{\rm fit}$ in a simulated galaxy with dispersion $\sigma_{\rm real} = 4$ km s^{-1} and different binary fractions B. The solid vertical line gives the error bar, equal to one standard deviation in $\sigma_{\rm fit}$ values calculated in 300 random realizations, which is similar for all points on the graph. The solid line uses a uniform prior in B, while the dashed line uses a prior $P(B)$ obtained from multi-epoch observations of the stars. The dot-dashed line is calculated from the biweight robust estimator. We adopted a measurement error of 2 km s^{-1} for all stars: (a) $\sigma_{\rm real} = 4$ km s^{-1}, 100 stars and (b) $\sigma_{\rm real} = 4$ km s^{-1}, 500 stars.



(a) $\sigma_{real} = 10$ km/s, 100 stars (b) $\sigma_{real} = 10$ km/s, 500 stars

Figure 14. Best-fit values for the dispersion $\sigma_{\rm fit}$ in a simulated galaxy with dispersion $\sigma_{\rm real} = 10$ km s^{-1} and different binary fractions B. Error bar gives one standard deviation in $\sigma_{\rm fit}$ values calculated in 300 random realizations, which is similar for all points on the graph. The solid line uses a uniform prior in B, while the dashed line uses a prior $P(B)$ obtained from multi-epoch observations of the stars. The dot-dashed line is calculated from the biweight robust estimator. We adopted a measurement error of 2 km s^{-1} for all stars: (a) $\sigma_{\rm real} = 10$ km s^{-1}, 100 stars and (b) $\sigma_{\rm real} = 10$ km s^{-1}, 500 stars.

the binary fraction. The most rigorous approach would use a joint likelihood $L(v_1, \Delta v|\Delta t)$, which can be generated from the Monte Carlo. This would have the advantage that individual binary stars which are inflating the dispersion *and* exhibit a large velocity change Δv would be recognized as such, and weighted accordingly. While this method may be necessary for obtaining constraints in small data sets of less than 100 stars, for larger samples we can adopt a simpler approach: first derive a posterior $P(B)$ in binary fraction by the multi-epoch analysis outlined in Section 4, and subsequently take $P(B)$ as a prior in B for the single-epoch analysis. The usefulness of this method depends critically on the size of the multi-epoch sample, as this determines the constraint on binary fraction (see Figure 3).

To evaluate this method, it is important to know how the best-fit dispersion $\sigma_{\rm fit}$ obtained by this procedure may differ from the intrinsic dispersion $\sigma_{\rm real}$ due to small number statistics. To investigate this, we analyze samples consisting of 100 and 500 stars and repeat the procedure over 300 randomly generated realizations of each sample. The range of $\sigma_{\rm fit}$ values obtained for galaxies with dispersions of 4 km s^{-1} and 10 km s^{-1} are plotted in Figures 13 and 14, respectively. Also included are the results if the prior in B is taken from multi-epoch observations of the same stars in the sample, in the manner outlined above. For comparison, we also plot the dispersion obtained from the biweight robust estimator, which is roughly comparable to the dispersion obtained by using a 3σ-clipping routine. This plot shows that the biweight is biased to large dispersions by an amount which depends on the binary fraction; furthermore, the bias for a 4 km s^{-1} dispersion is much larger compared to the error bars than for a 10 km s^{-1} dispersion. By comparison,

the best-fit dispersion $\sigma_{\rm fit}$ using a uniform prior in B is biased by a smaller amount which depends on the true binary fraction, up to 10%. However, if the prior $P(B)$ is calculated from multi-epoch samples of as few as 100 stars, the bias is almost entirely removed. Thus, even a fairly weak constraint on the binary fraction is sufficient to remove most of the bias.

Unfortunately, the likelihood analysis outlined above is of limited utility in actual data sets because of contamination by foreground Milky Way stars. The usual criteria to determine membership of a star include its location on the color–magnitude diagram, metallicity, and radial velocity. Outliers in the distribution of these variables are either excluded, or else weighted by a low membership probability assigned according to a specific algorithm (cf. Walker et al. 2009b). However, among the stars with single-epoch measurements, only those stars which lie on the high-velocity tail of the velocity distribution can help constrain the binary fraction. If even a small number of high-velocity binary stars are excluded from the data or weighted by a low membership probability, the inferred binary fraction will be biased to low values.

In principle, this problem can be resolved by including the velocity distribution of foreground Milky Way stars in the likelihood analysis, obtained by a kinematic model of the disk and bulge components (Robin et al. 2003). The success of this method will depend critically on the degree of overlap between the two distributions, which is partly determined by the systemic velocity of the background galaxy. It is also essential that binarity in the Milky Way is accounted for, since this adds a substantial high-velocity tail to the foreground velocity distribution. Accounting for binarity in the likelihoods of both the foreground and background stars may also lead to improved membership probabilities when combined with the expectation maximization algorithm of Walker et al. (2009b). Including both binarity and foreground Milky Way stars in a likelihood analysis is certainly of interest for obtaining the best possible constraints from single-epoch velocity data, but is beyond the scope of this work.

7. CONSTRAINING THE DISTRIBUTION OF PERIODS

Here, we address the question of what form the distribution of periods might take in regions beyond the solar neighborhood and whether this can be constrained by radial velocity data. Simulations of star formation (Machida et al. 2009; Tohline 2002; Bate 2000) indicate that the statistical properties of binary systems are determined during star formation via turbulent fragmentation of a rotating gas cloud. The distributions in the orbital parameters generally undergo little subsequent modification by collisional processes, except in the high-density regions found at the centers of globular clusters (Hut et al. 1992; Pryor et al. 1988). This suggests that the distribution of periods in dwarf spheroidals and dwarf irregular galaxies may be of a similar form to that found in the solar neighborhood. We shall assume the log-normal is a fair approximation to the period distribution; it is reasonable to ask, however, to what extent its mean $\mu_{\log P}$ and dispersion $\sigma_{\log P}$ may be expected to differ from that of the solar neighborhood. Because of the difficulty of simulating binary star formation, at present we have an incomplete picture of how these parameters might vary depending on the star formation history of each galaxy. However, the following points can be made.

A semi-empirical model of isolated binary star formation by Fisher (2004) yielded values of $\sigma_{\log P}$ within the range 1.6–2.1, depending on the star formation efficiency of the initial gas cloud. Observations of pre-main-sequence stars in Milky Way stellar associations also show more peaked distributions than $\sigma_{\log P} = 2.3$ (Brandner & Koehler 1998). This suggests that the wider distribution observed in the solar neighborhood may be formed by a superposition of more sharply peaked binary distributions resulting from various star-forming environments. Accordingly, to be conservative in this paper we have considered values of $\sigma_{\log P} \in (0.5, 4)$. Likewise we have considered values of $\mu_{\log P} \in (-1, 4)$ (P in years). The low-$\mu_{\log P}$, high-$\sigma_{\log P}$ boundaries of these intervals describe velocity distributions with a highly distorted, non-Gaussian shape and therefore can be considered unlikely.

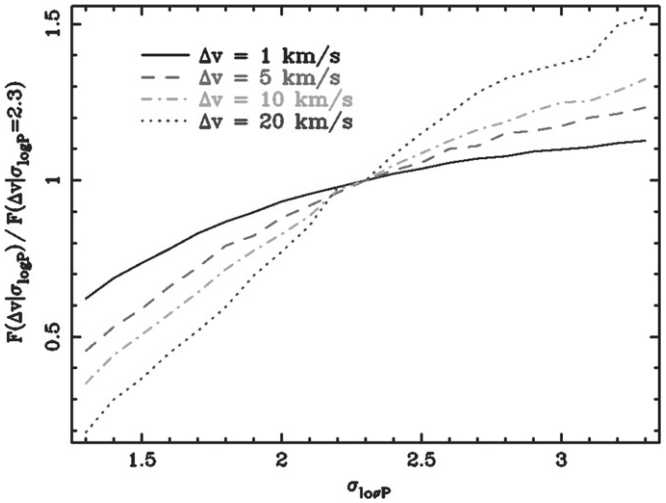


Figure 15. Behavior of the threshold fraction $F(\Delta v)$ if the width of the period distribution $\sigma_{\log P}$ (P in years) is varied, plotted for different velocity thresholds Δv. The y-axis is given by $f = F/F_0$, where F_0 is the fiducial threshold fraction assuming the value $\sigma_{\log P} = 2.3$ observed in the solar neighborhood (Duquennoy & Mayor 1991).

Supposing that $\mu_{\log P}$ and $\sigma_{\log P}$ may vary by the amount suggested by Fisher (2004), can these parameters be estimated independently of binary fraction? To answer this question, we reconsider the threshold fraction $F(\Delta v|\Delta t)$, defined as the fraction of stars with an observed change in velocity greater than a threshold Δv after a time Δt between measurements. For each threshold, let us define the ratio $f = F/F_0$ where F_0 is the fiducial threshold fraction obtained by setting $\sigma_{\log P} = 2.3$, $\mu_{\log P} = 2.24$. Given a time interval of one year and picking several thresholds Δv, we plot f as a function of $\sigma_{\log P}$ in Figure 15 and $\mu_{\log P}$ in Figure 16. The similarity in slope among the different curves gives a measure of degeneracy; if the curves were identical, the parameters would be completely degenerate with the binary fraction. However, if the difference between the curves is smaller than the statistical error bars, the parameters are effectively degenerate with B and we cannot hope to distinguish between them. The error bars on f at a particular Δv are given by (compare Equation (B3) in Appendix B):

$$\epsilon_f \approx \frac{2f}{\sqrt{N\bar{F}(\Delta v_{\rm tail}|\Delta t)}}, \tag{24}$$

where ϵ_f is the 95% confidence limit in f, N is the number of stars, and $\Delta v_{\rm tail}$ is defined as before (for a rough approximation, given a measurement error σ_m one can take $\Delta v_{\rm tail} \approx 2\sigma_{2e}$ where $\sigma_{2e} \approx \sqrt{2}\sigma_m$). For $\Delta t = 1$ yr and a measurement error of 2 km s^{-1}, Figure 4 gives $\bar{F}(\Delta v_{\rm tail}) \approx 0.1 \times B$. If $B \approx 0.5$, this gives $\epsilon_f \approx 9f/\sqrt{N}$. Assuming a sample of $N = 100$

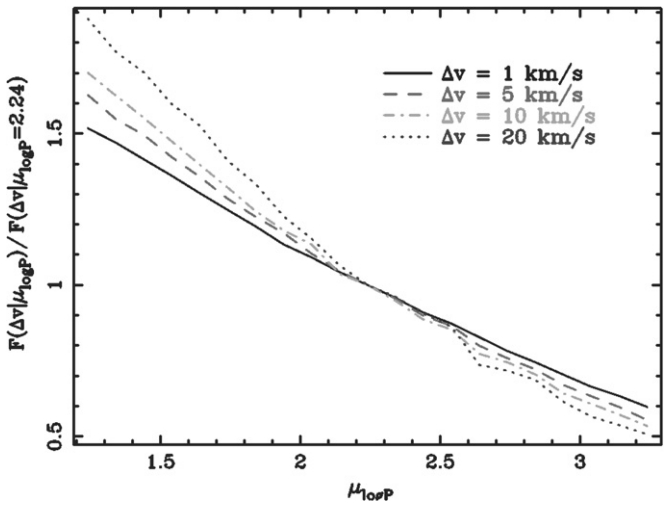


Figure 16. Behavior of the threshold fraction $F(\Delta v)$ if the mean of the period distribution $\mu_{\log P}$ (P in years) is varied, plotted for different velocity thresholds Δv. The y-axis is given by $f = F/F_0$, where F_0 is the fiducial threshold fraction assuming the value $\mu_{\log P} = 2.24$ observed in the solar neighborhood (Duquennoy & Mayor 1991).

stars, we have $\epsilon_f \approx 0.9$, far too large to distinguish between the different curves in Figures 15 and 16. Thus to within the statistical error in a sample of a few hundred stars, $\sigma_{\log P}$ and $\mu_{\log P}$ are degenerate with the binary fraction over nearly the entire range of Figures 15 and 16; only with $N \gtrsim 1000$ are the error bars small enough to break this degeneracy. We conclude that multi-epoch samples of ≈1000 stars or more are required in order to constrain the period distribution of a binary population by the method proposed here.

8. CONCLUSION

We have demonstrated a procedure to estimate the intrinsic velocity dispersion of dwarf spheroidal galaxies more precisely than in previous studies by accounting for the effect of binary orbital motion. By measuring the threshold fraction of a stellar sample (defined in Section 3.2), the correction to the velocity dispersion due to binary motion can be estimated; fitting functions are provided for this method (Equations (16), (17)–(19)). We have also demonstrated a method to estimate the threshold fraction, together with other properties of the binary population, more rigorously from multi-epoch data via a Bayesian or maximum likelihood approach. We conclude with the following points.

1. The measured velocity dispersions of dwarf spheroidal galaxies between 4 and 10 km s^{-1} are unlikely to be inflated by more than 30% by binaries (Section 5.6); if the sample consists only of red giants, the upper limit becomes 20%. This conclusion can be made with confidence because—as we showed in Section 5—the correction to the dispersion holds independent of the model used to describe the binary population, provided that the dispersion is calculated using a high-velocity cutoff (e.g., by a 3σ-clipping routine). This is fortunate since the binary fraction and other model parameters used to describe the binary population are very difficult to constrain independently of each other with samples of only a few hundred stars (Section 7).
2. With a measurement error on the order of 1 km s^{-1}, we find that 1–2 yr is an optimal interval between measurements for measuring the threshold fraction, since the fraction of stars with a measurable change in velocity does not rise significantly when the interval is extended beyond two years (Figure 4). Furthermore, to constrain the binary fraction and other properties of the binary population, it is a more profitable strategy to make two-epoch measurements over a larger sample of stars, as opposed to adding more repeat measurements to an already existing two-epoch sample. This is necessary to overcome the large scatter in the binary fraction of samples with less than a few hundred stars. We also find that multi-epoch samples of more than 1000 stars would be required in order to constrain the distribution of periods independently of the binary fraction by the method proposed in this paper.
3. While the binary population can in principle be constrained by single-epoch data, in practice this is very difficult because of contamination by non-member stars. Since an outlier in the velocity distribution cannot be verified as a binary star without multiple epoch measurements, it may be erroneously labeled a non-member star and excluded from the data sample (or weighted by a low membership probability). This would bias the estimated binary fraction to low values and result in an inflated dispersion estimate. Even in single-epoch samples, however, this problem may be overcome by including a likelihood for the foreground Milky Way stars in a Bayesian analysis. This analysis can also be combined with multi-epoch data to provide better constraints, and in principle would extract the best constraints from both single- and multi-epoch velocity data.

We thank Erik Tollerud for providing valuable feedback and many illuminating discussions throughout the course of this project. This work was supported in part by NSF grant AST-0607746 and NASA grant NNX09AD09G.

APPENDIX A

CALCULATING THE THRESHOLD FRACTION FOR DIFFERENT MEASUREMENT ERRORS

In Section 5, we showed how to correct the dispersion (Equation (16)) by using the threshold fraction F, defined as the fraction of stars with an observed change in velocity greater than a threshold Δv after a time Δt between measurements. If the threshold fraction is calculated directly from the data, then measurement error must be considered in Equation (16). While the measured dispersion can be easily corrected for measurement error according to $\sigma^2_{\rm meas} = \sigma^2 - \sigma^2_m$ (where $\sigma_{\rm meas}$ is the measured dispersion and σ_m is the measurement error), correcting the threshold fraction is less trivial. To correct the threshold fraction for measurement error, first we note that the degeneracy of the period parameters with binary fraction means that the velocity distribution can be approximately split into two parts: a small-v part which acts effectively like a δ-function similar to that in Equation (10), and a large-v tail which scales linearly with B, $\mu_{\log P}$, and $\sigma_{\log P}$. Exactly where to do the "splitting" is dictated largely by the measurement error, as the tail becomes prominent only at velocities beyond $\Delta v \approx \sigma_{2e}$. We therefore split the function at $\Delta v_{\rm tail} = \gamma\sigma_{2e}$, where γ is a proportionality constant with a very weak (if any) dependence on σ_{2e}. The threshold fraction can then be written approximately as follows:

$$F(\Delta v|\Delta t, B, \sigma_{2e}) \approx (1-\mathcal{N})\mathrm{erfc}\left[\frac{\Delta v}{\sqrt{2}\sigma_{2e}}\right] + BJ(\Delta v|\sigma_{2e}), \quad \text{(A1)}$$

where

$$J(\Delta v|\sigma_{2e}) = \int_{-\infty}^{\infty} \mathrm{erfc}\left[\frac{\Delta v - \Delta v'}{\sqrt{2}\sigma_{2e}}\right] \frac{g_b(\log|\Delta v'|)}{|\Delta v'|\ln 10} d(\Delta v'), \tag{A2}$$

$$\mathcal{N} = B\,J(\gamma\sigma_{2e}|\sigma_{2e}). \tag{A3}$$

The normalization factor $\mathcal{N}$ here has replaced B in Equation (14) and varies linearly with B, $\mu_{\log P}$, and $\sigma_{\log P}$. Now as long as Δv is approximately equal to or larger than σ_{2e}, the error function in the integrand of Equation (A2) is approximately a smoothed step function. For $\Delta v/\sigma_{2e} \gtrsim 1$ we can therefore make the approximation

$$J(\Delta v|\sigma_{2e}) \approx \alpha\left(\frac{\Delta v}{\sigma_{2e}}\right) F_0(\Delta v), \tag{A4}$$

where $F_0(\Delta v) \equiv F(\Delta v|\sigma_{2e} = 0)$ denotes the threshold fraction without measurement error. Note that in the limit $\sigma_{2e} \to 0$ ($\Delta v/\sigma_{2e} \to \infty$), the integrand of Equation (A2) becomes a step function so that $\alpha \to 1$. On the other end, as $\Delta v/\sigma_{2e}$ is made smaller, the integral in Equation (A2) includes more of the central peak so α becomes larger. Using Equation (A4), we can write $\mathcal{N} = \kappa F_0(\Delta v)$ where

$$\kappa = \frac{\alpha(\gamma)F_0(\gamma\sigma_{2e})}{\alpha\left(\frac{\Delta v}{\sigma_{2e}}\right)F_0(\Delta v)}. \tag{A5}$$

Since F_0 is the tail of a log-normal distribution, over the scale of several km s^{-1} we have $F_0(\Delta v) \approx (\Delta v)^{-n}$, where n is an exponent close to 1. Therefore κ can be written as a function of $\Delta v/\sigma_{2e}$:

$$\kappa\left(\frac{\Delta v}{\sigma_{2e}}\right) \approx \frac{\gamma^{-n}\alpha(\gamma)}{\alpha\left(\frac{\Delta v}{\sigma_{2e}}\right)}\left(\frac{\Delta v}{\sigma_{2e}}\right)^n. \tag{A6}$$

Substituting this result into Equation (A1), we have

$$F(\Delta v|\sigma_{2e}) \approx \mathrm{erfc}\left[\frac{\Delta v}{\sqrt{2}\sigma_{2e}}\right] + \beta\left(\frac{\Delta v}{\sigma_{2e}}\right) F_0(\Delta v), \tag{A7}$$

where

$$\beta\left(\frac{\Delta v}{\sigma_{2e}}\right) \equiv \alpha\left(\frac{\Delta v}{\sigma_{2e}}\right)\left\{1 - \kappa\left(\frac{\Delta v}{\sigma_{2e}}\right)\mathrm{erfc}\left[\frac{\Delta v}{\sqrt{2}\sigma_{2e}}\right]\right\}. \tag{A8}$$

Equation (A7) is the key result: it means that the threshold fraction F with a given measurement error is related to the measurement error-free value F_0 by a linear transformation, and the only extra information required to make this transformation is the function $\beta\left(\frac{\Delta v}{\sigma_{2e}}\right)$. As a check, taking the limit $\sigma_{2e} \to 0$ we have $\alpha \to 1$ and $\kappa \to 0$, so that $\beta \to 1$ as it should. We use the Monte Carlo simulation to plot β as a function of $\Delta v/\sigma_{2e}$ for different velocity thresholds and find the function β is nearly the same regardless of threshold, which justifies the approximations taken to reach Equation (21). We find that for $\Delta v/\sigma_{2e} \gtrsim 1$, the parameter κ shows only slight variation over the range for which the error function is non-negligible. We also find that α is well approximated by an exponential, so that β is well fit by the following function:

$$\beta(x) = \left(1 + ae^{-\frac{x}{x_s}}\right)\left\{1 - \bar{\kappa}\cdot\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)\right\}, \tag{A9}$$

where $x = \Delta v/\sigma_{2e}$ and the best-fit parameters are $a = 0.05$, $x_s = 5$, and $\bar{\kappa} = 1.3$.

APPENDIX B

FORMULA FOR NUMBER OF STARS REQUIRED TO CONSTRAIN THE BINARY FRACTION

For a set of N stellar velocities (with $N > 100$) measured at two different epochs separated by time Δt, consider the number of stars n with change in velocity greater than some threshold value Δv. If n were to be measured over many random realizations drawn from a particular galaxy, it would follow a Poisson distribution. For the time being, let us take the measurement error σ_{2e} to be zero. The mean number $\bar{n}$ is related to the mean threshold fraction of binaries $\bar{F}_b(\Delta v|\Delta t, \sigma_{2e} = 0)$ by $\bar{n} = NB\bar{F}_b$ where N is the total number of stars. If we pick Δv small enough such that this number is larger than 10, the Poisson distribution is approximately Gaussian with standard deviation $\sigma_n \approx \bar{n} \approx \sqrt{NB\bar{F}_b}$. Since the fraction of stars is $F = n/N$, we obtain the error in the measured threshold fraction F:

$$\sigma_F = \sqrt{\frac{\bar{F}(\Delta v|\Delta t, B, \sigma_{2e} = 0)}{N}}, \tag{B1}$$

where we also used the relation $\bar{F} = B\bar{F}_b$. Now if σ_{2e} is nonzero, the smallest value of Δv which is largely unaffected by the measurement error will occur roughly at the point $\Delta v_{\rm tail}$ where $\bar{F}(\Delta v_{\rm tail}|\Delta t, B, \sigma_{2e} = 0) = \bar{F}(\Delta v_{\rm tail}|\Delta t, B = 0, \sigma_{2e})$; this is where the "tail" in the distribution due to binaries begins to dominate. Therefore, we pick this point as giving the best constraint on B. (For a rough approximation, one can also use $\Delta v_{\rm tail} \approx 2\sigma_{2e}$.) If we measure the fraction of stars with change in velocity greater than $\Delta v_{\rm tail}$, given by $F = n/N$, our "best-fit" binary fraction b is then defined by $F = b\bar{F}_b(\Delta v_{\rm tail}|\Delta t, B, \sigma_{2e} = 0)$. Combining this with $\bar{F} = B\bar{F}_b$ and Equation (B1) gives the standard deviation of the best-fit binary fraction:

$$\sigma_b = \sqrt{\frac{B}{N \times \bar{F}_b(\Delta v_{\rm tail}|\Delta t, \sigma_{2e} = 0)}}. \tag{B2}$$

Now calling the 95% confidence limit $\epsilon_b = 2\sigma_b$ and solving for N yields Equation (15). The approximation becomes less accurate as B tends toward very small (close to 0) and large (close to 1) values. If B is sufficiently close to 0 such that n is less than 10, the Gaussian limit no longer holds; in that limit, the uncertainty will be larger than that given in Equation (B2). If B is close to 1, the width of the Gaussian is larger than the true uncertainty since B-values greater than 1 are not allowed. In Figure 3, the approximation formula is graphed as a function of N for several values of B and compared to the 95% confidence interval of the posterior $P(B)$ obtained from a Bayesian analysis of the simulated data, averaged over a hundred realizations.

Using the best-fit binary fraction obtained by the threshold fraction at $\Delta v_{\rm tail}$, we can find the standard deviation of the best-fit threshold fraction at a given threshold Δv by substituting the relations $F = b\bar{F}_b(\Delta v|\Delta t, B, \sigma_{2e} = 0)$ and $\bar{F} = B\bar{F}_b$ into Equation (B2), with the result

$$\sigma_{F,\rm fit} = \frac{\bar{F}(\Delta v|\Delta t, \sigma_{2e} = 0)}{\sqrt{N\bar{F}(\Delta v_{\rm tail}|\Delta t, \sigma_{2e} = 0)}}. \tag{B3}$$

This equation differs from Equation (B1) in that it uses the best-fit binary fraction to infer the threshold fraction at thresholds

$\Delta v > \Delta v_{\rm tail}$. Since the scatter in F is smaller at $\Delta v_{\rm tail}$, this leads to a better constraint than if the threshold fraction is measured directly. Comparing Equations (B3) and (B1) and using the fact that the threshold fraction F decreases monotonically in Δv, we see that the error $\sigma_{F,\rm fit} < \sigma_F$ as expected. Equation (B3) approximates the error in the threshold fraction estimated by the Bayesian approach outlined in Section 4. We caution, however, that if one uses high thresholds ($\Delta v > 10$ km s^{-1}) where there are few data points, the inferred binary fraction may be incorrect as the likelihood approach may incorrectly extrapolate from the distribution at lower velocities.

REFERENCES

Bate, M. R. 2000, MNRAS, 314, 33
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Brandner, W., & Koehler, R. 1998, ApJ, 499, L79
Church, R. P., Dischler, J., Davies, M. B., Tout, C. A., Adams, T., & Beer, M. E. 2009, MNRAS, 395, 1127
Duquennoy, A., & Mayor, M. 1991, A&A, 248, 485
Eggleton, P. P. 1983, ApJ, 268, 368
Fischer, D. A., & Marcy, G. W. 1992, ApJ, 396, 178
Fisher, R. T. 2004, ApJ, 600, 769
Geha, M., Willman, B., Simon, J. D., Strigari, L. E., Kirby, E. N., Law, D. R., & Strader, J. 2009, ApJ, 692, 1464
Girardi, L., Grebel, E. K., Odenkirchen, M., & Chiosi, C. 2004, A&A, 422, 205
Goldberg, D., Mazeh, T., & Latham, D. W. 2003, ApJ, 591, 397
Goodwin, S. P., Kroupa, P., Goodman, A., & Burkert, A. 2007, in Protostars and Planets V, The Fragmentation of Cores and the Initial Binary Population, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 133
Halbwachs, J., Mayor, M., Udry, S., & Arenou, F. 2004, RevMexAA Conf. Ser., 21, 20
Hargreaves, J. C., Gilmore, G., & Annan, J. D. 1996, MNRAS, 279, 108
Hut, P., et al. 1992, PASP, 104, 981
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2002, in ASP Conf. Ser. 285, Modes of Star Formation and the Origin of Field Populations, ed. E. K. Grebel & W. Brandner (San Francisco, CA: ASP), 86
Kroupa, P., Tout, C. A., & Gilmore, G. 1990, MNRAS, 244, 76
Łokas, E. L., Mamon, G. A., & Prada, F. 2005, MNRAS, 363, 918
Machida, M. N., Omukai, K., Matsumoto, T., & Inutsuka, S.-i. 2009, arXiv:0907.3257
Mateo, M., Olszewski, E., Welch, D. L., Fischer, P., & Kunkel, W. 1991, AJ, 102, 914
Mazeh, T., Goldberg, D., Duquennoy, A., & Mayor, M. 1992, ApJ, 401, 265
Olszewski, E. W., Pryor, C., & Armandroff, T. E. 1996, AJ, 111, 750
Paczyński, B. 1971, ARA&A, 9, 183
Pryor, C. P., Latham, D. W., & Hazen, M. L. 1988, AJ, 96, 123
Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523
Salpeter, E. E. 1955, ApJ, 121, 161
Simon, J. D., & Geha, M. 2007, ApJ, 670, 313
Tohline, J. E. 2002, ARA&A, 40, 349
Tonry, J., & Davis, M. 1979, AJ, 84, 1511
Trimble, V. 1990, MNRAS, 242, 79
Walker, M. G., Mateo, M., & Olszewski, E. W. 2009a, AJ, 137, 3100
Walker, M. G., Mateo, M., Olszewski, E. W., Sen, B., & Woodroofe, M. 2009b, AJ, 137, 3109
Willman, B., et al. 2005, ApJ, 626, L85
Zucker, D. B., et al. 2006, ApJ, 650, L41